AD_____

Award Number: W81XWH-05-1-0204

TITLE: Identification, Characterisation and Clinical Development of the New General of Breast Cancer Susceptibility Alleles

PRINCIPAL INVESTIGATOR: Nazneen Rahman, M.D., Ph.D.

CONTRACTING ORGANIZATION: The Institute of Cancer Research
London SW7 3RP; United Kingdom

REPORT DATE: March 2009

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| 1. REPORT DATE | 2. REPORT TYPE | 3. DATES COVERED |
|---|---|---|
| 01-03-2009 | Annual | 1 Mar 2008 - 28 Feb 2009 |

**4. TITLE AND SUBTITLE**
**Identification, Characterisation and Clinical Development of the New Generation of Breast Cancer Susceptibility Alleles**

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**
**W81XWH-05-1-0204**

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**
Nazneen Rahman, M.D. Ph.D.

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**

The Institute of Cancer Research
London SW7 3RP, United Kingdom

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY**
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION / AVAILABILITY STATEMENT**

Approved for public release; distribution unlimited

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**
Breast cancer is a common disease in women but the causes are still largely unknown. There is considerable evidence that genetic factors play an important role in causing breast cancer, but the genes involved in the majority of breast cancers are currently unknown. Our aim is to identify genetic factors that increase the risk of breast cancer occurring. We have collected samples and clinical information from over 4000 breast cancer families. We compare the frequency of genetic factors in these cases with control individuals. Over the last year we have been engaged in two complementary strategies. 1) Undertaking genome-wide association analyses to identify common, low-penetrance variants that increase breast cancer risk by a modest amount. Our collaborative endeavours in this area have already led to the identification of several variants and we are currently undertaking the largest study to date that involves 4000 cases and 6000 controls. 2) Candidate gene analysis by sequencing of DNA repair genes for rare, intermediate-penetrance variants. This previously led to our identification of 4 breast cancer genes, *CHEK2*, *ATM*, *PALB2* and *BRIP1*. We have also been investigating the interaction between breast cancer genes and have demonstrated a negative interaction between the intermediate-penetrance and high-penetrance genes, presumably because they operate in the same pathways.

**15. SUBJECT TERMS**

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | UU | 42 | USAMRMC |
| U | U | U | | | 19b. TELEPHONE NUMBER *(include area code)* |

**Standard Form 298 (Rev. 8-98)**
**Prescribed by ANSI Std. Z39.18**

## Table of Contents

**Introduction**

Breast cancer is a common disease in women but the causes are still largely unknown. There is considerable evidence to suggest that genetic factors play an important role in causing breast cancer. In recent years, our understanding of genetic predisposition to breast cancer has advanced significantly. Three classes of predisposition factor, categorised by their associated risks of breast cancer, are currently known. *BRCA1* and *BRCA2* are high-penetrance breast cancer predisposition genes identified by genome-wide linkage analysis and positional cloning. Mutational screening of genes functionally related to *BRCA1* and/or *BRCA2* has revealed four genes, *CHEK2, ATM, BRIP1* and *PALB2*, mutations of which are rare and confer an intermediate risk of breast cancer. Association studies have, thus far, identified ten common variants associated with low-penetrance breast cancer predisposition (*1,2*). Our group has had a major role in the elucidation of breast cancer predisposition identifying *BRCA2*, *CHEK2*, *BRIP1*, *PALB2*, clarifying the role of *ATM* in breast cancer predisposition and contributing to the international collaborative efforts to identify the common, low-penetrance breast cancer alleles (*3-8*). However, despite these discoveries, most of the familial risk of breast cancer remains unexplained.

Our aim therefore remains the identification and characterisation of the genetic factors that increase the chance of breast cancer occurring. We have collected clinical information and samples from over 4000 breast cancer families to facilitate these aims. Over the reporting period we have been engaged to three main areas, outlined in further detail below 1) Undertaking the largest genome-wide association study (GWAS) in breast cancer to date, which includes 4000 familial cases and 6000 controls in the first phase and 2) Continuing with our candidate-gene sequencing analyses of DNA repair genes which previously led to the identification of *CHEK2*, *ATM*, *BRIP1* and *PALB2*. 3) Investigating the interactions between

mutations in rare, intermediate-penetrance breast cancer predisposition genes and *BRCA1/BRCA2* mutations. An important focus for us currently, is the evaluation and optimization of next-generation sequencing technologies to allow us to take the candidate gene sequencing to a genome-wide level, which we hope to begin within the next reporting period.

**Body**

As part of the programme of work we defined five tasks. The progress towards the tasks is outlined in detail below. As genetics is a fast-moving field and there have been considerable technological advances over the last few years, we have revised some aspects of the original tasks to ensure that we deliver our aims in the most effective and efficient manner possible (see below.

*Task 1: Evaluate the contribution of BRCA1 and BRCA2 exonic deletions and duplications to breast cancer susceptibility.*

This task is complete – see previous report.

*Task 2. Perform familial case-control analyses of non-synonymous coding single nucleotide polymorphisms (SNPs) in DNA repair genes in familial breast cancer cases.*

This task is complete – see previous report and paper (*9*)

*Task 3. Characterise the histopathology and immunohistochemistry of familial breast cancer.*
We have amended this task slightly to reflect the new insights that have emerged. We are undertaking and have achieved the following in this reporting period

a) *Accruing data on the ER status of the 4000 cases included in the genome-wide association study so that we can analyse the data by ER status.*

To date we have collected this information on 678 cases. We are also collecting data on PR and HER2 and other pathological characteristics where available. However, initial GWAS studies have demonstrated that ER status is the strongest phenotypic surrogate (*10*). Therefore we have prioritized obtaining ER status on as many cases as possible for the primary analysis and plan to go back and undertake a more detailed review, and ideally construct tissue microarrays, at a later date.

b) *Collection and analysis of sporadic breast cancer cases with triple-negative tumors (ER, PR and HER2 negative) which we are stratifying by BRCA1/2 status.*

Genetic and biological data indicate that triple-negative, basal-like tumors are a distinctive sub-phenotype of breast cancers that may have different underlying causes (*10*). There is a known strong association of triple-negative tumor phenotype and *BRCA1* mutations. However, the contribution of *BRCA1* to triple-negative breast cancer in the absence of a strong family history remains unclear and is a source of considerable confusion diagnostically. We aim to address this question by screening substantial (at least 300) individuals with triple-negative tumors for *BRCA1* mutations. Over the last year we have collected over 150 cases and we anticipate that we will undertake this experiment within the next reporting period. The *BRCA*-negative triple-negative cases will be incorporated into our ongoing research to identify new genes.

c) *Tumor collection and pathological, immunohistochemical and loss of heterozygosity analyses to define the tumor characteristics associated with the rare, intermediate-penetrance breast cancer susceptibility genes, ATM, BRIP1, CHEK2, PALB2.*

To date we have 90 cases due to *CHEK2* mutations, 19 with *ATM*, 19 with *BRIP1* and 17 with *PALB2* mutations. We are continuing to screen these genes in our breast cancer families to identify additional cases. We have obtained tumor information and/or data from half of the cases. We plan to try to obtain as many tumor samples as possible during 2009 and will then undertake the loss of heterozygosity analyses and detailed path review of all collected cases in 2010 with a view to submitting the results for publication towards the end of 2010.

*Task 4. Perform genome-wide familial case-control analyses of non-synonymous coding SNPs,*

This task is completed (as originally proposed) and published – see previous report and paper (*9*). We undertook a direct analysis of the known (at that time) 15,000 non-synonymous coding SNPs (*9*). Our current genome-wide association study in 4000 cases and 6000 controls will additionally allows us to tag any non-synonymous coding SNPs of 5% or greater frequency (see below) and will therefore will allow us to interrogate the remaining non-synonymous coding SNPs. However, our understanding of this area has changed considerably within the last few years and we have thus extended and revised this experiment. Initially, it was considered most plausible that common variation underpinning common disease would most likely reside in genes and that non-synonymous variation would be the most likely causative alleles. However, in breast cancer (and indeed many other common diseases) very few of the common low-penetrance alleles have been in genes and indeed some are in gene deserts (2, 8). Hence, our ongoing efforts to evaluate common variation in breast cancer susceptibility are encompassing the full genome and are not focusing solely on coding variants (see below).

*Task 5. Identify and characterize low-penetrance breast cancer susceptibility alleles*

The aim of this task remains the same. However, we have adapted the original task to capitalize on the results of our (and other researchers) research as well as the considerable technological advances that have occurred over the last few years.

a) *Undertake a second-generation genome-wide association study in 4000 cases and 6000 controls to identify common, low-penetrance breast cancer susceptibility alleles.*

Genome-wide association studies (GWAS) have already identified novel susceptibility loci for several common diseases. In 2007, we completed the first such study in breast cancer in collaboration with Professor Doug Easton. This was based on 400 genetically enriched breast cancer cases and 400 controls typed for over 220,000 SNPs. These SNPs were correlated with ~71% of known common SNPs, at $r^2$>0.5. Putative associations were followed up in ~26,000 cases and 26,000 controls. This study provided clear evidence for five novel breast cancer susceptibility loci (8). Further studies of this initial scan have recently led to the identification of two further variants on 3p24 and 17q23 (*11*). A further 3 loci have been identified (Hunter et al, 2007; Stacey et al, 2007). Together these 10 loci explain about 6% of the familial risk of the disease, so that in total approximately 25% of the familial risk is now explained (reviewed in (*2*)).

Although the GWAS studies undertaken to date have been successful they were underpowered as the genome-wide phase was undertaken in relatively small series. Last year, on the basis of our previous data we successfully applied for a grant to undertake a second-generation GWAS scan which is the largest undertaken to date by some considerable margin. We have genotyped 4000 familial breast cancer cases on the Illumina 670 SNP chip and will be comparing this to genome-wide data from 6000 control individuals. The grant funded the genome-wide genotyping of cases but the pre-genotyping curation, statistical analyses and replication will be undertaken as part of our EOH initiatives. During this reporting period we spent a very considerable amount of time preparing and curating the samples prior to the genotyping. We sent the samples in

November and we have just received the data. We will be undertaking the statistical analyses both for standard SNP associations and for copy number variants (CNVs) and replication of positive associations over the next 6-12 months. We will replicate the top 30 SNPs in-house in 3500 additional cases and 3500 controls using Taqman. We will also then contribute the genome-wide data to an international meta-analysis that will hopefully include data from all the breast cancer GWAS scans performed to date.

The power calculations for the experiment suggest that our GWAS scan will provide better than 80% power to detect any locus conferring a per allele relative risk of 1.1, provided the allele frequency is at least 10%, and any allele with a frequency of 5% conferring a relative risk of 1.2. The calculations assume the expected enrichment of the stage 1 cases by virtue of their family history, and are therefore assume that interactions between loci are approximately multiplicative. The power is principally dependent on the familial relative risk conferred by the SNP, and power is better than 90% to detect any locus explaining at least 0.2% of the familial risk.

b) *Undertake case-control resequencing of candidate DNA repair genes to identify further rare, low-penetrance genes.*

In the last year we have particularly focused on evaluating the DNA repair gene, *GEN1*. GEN1 is a recently identified Holliday junction resolvase with an important role in recombinational repair (*14*). As such it is potentially important in double-strand break repair via the homologous recombination pathway. The breast cancer predisposition genes *BRCA1*, *BRCA2*, *ATM*, *TP53*, *CHEK2*, *BRIP1* and *PALB2* are all involved in these pathways (reviewed in Turnbull and Rahman, 2008). In addition, somatic truncating *GEN1* mutations in two breast cancer cell lines have been reported (*15*). Taken together, this recommended *GEN1* as a plausible candidate breast cancer predisposition gene. To investigate this we screened the full coding sequence

and intron-exon boundaries of *GEN1* in constitutional DNA from 192 familial breast cancer cases (that are negative for mutations in *BRCA1* and *BRCA2*) and 192 controls.

We identified a five base pair deletion, c.2515_2519delAAGTT. We extended analysis of this truncation to 540 cases and controls. The mutation was present in 47/1072 case chromosomes and 47/1050 control chromosomes successfully analysed. Hence there was no evidence that this variant conferred an excess risk of breast cancer (P=0.9). The mutation is predicted to result in a protein product lacking the terminal 69 amino acids of GEN1 and we confirmed this in cDNA analyses. Furthermore, we identified case and control individuals homozygous for the truncating mutation, indicating that the last 69 amino acids of GEN1 are dispensable for its function.

We also identified 7 synonymous and 12 non-synonymous variants. None were predicted to be pathogenic by *in silico* analyses. We detected 14 of these variants at similar frequencies in cases and controls with four being common in both groups. Three rare non-synonymous variants were found in cases but not controls and two rare non-synonymous variants were found in controls but not cases.

These data indicate that, despite the crucial role of GEN1 in recombinational repair and the identification of somatic *GEN1* mutations in breast cancer cell lines, *GEN1* is not a breast cancer predisposition gene. It also highlights the importance of undertaking suitable genetic experiments, including full analysis of a candidate gene in controls, when evaluating the potential role of a gene in conferring susceptibility to disease. We are preparing these data for publication.

c) *Evaluate novel breast cancer susceptibility alleles in BRCA1 / BRCA2 families to determine whether they modify or interact with these genes in breast cancer.*

The issue of epistatis between breast cancer susceptibility alleles is a topical and important area, particularly as the recently identified variants *individually* confer very small increases in risk. Preliminary analyses from genome-wide association studies suggest that the common, low-penetrance breast cancer susceptibility alleles act multiplicatively with each other (*8*). However, association studies performed within *BRCA1* and *BRCA2* mutation-positive families suggest that at least some of these SNPs confer additional risk in the presence of *BRCA2* but not *BRCA1* mutations, although this may primarily reflect the association of the SNPs with ER positive tumors (as *BRCA1* is typically associated with ER negative tumors) (*16,10*).

We have been evaluating the interaction of the rare, intermediate-penetrance breast cancer genes with *BRCA1* and *BRCA2*. Such studies are challenging because of the rarity of mutations and the requirement to screen substantial numbers of cases through the complete coding sequence to identify them. Within the last year we have completed our analyses of *ATM* and *CHEK2* in *BRCA1* and *BRCA2* mutation-positive individuals.

We identified *CHEK2* mutations in 90/4741 familial *BRCA1/2*-negative breast cancer families compared to 8/2045 controls (P=2.9x10-7) and 1/345 *BRCA1*-positive families (P=0.01) and 1/300 *BRCA2*-positive families P=0.02). Similarly we identified *ATM* mutations in 19/1436 familial *BRCA1/2* negative breast cancer families compared to 0/273 *BRCA1*-positive families (P=0.01) and 1/335 *BRCA2*-positive families (P=0.02).

These data confirm and extend our initial data identifying *ATM* and *CHEK2* as breast cancer predisposition genes and demonstrates a negative interaction with *BRCA1* and *BRCA2*. i.e. mutations in *ATM* and *CHEK2* do not confer an increased risk of breast cancer on the

background of *BRCA1* or *BRCA2* mutations. This is plausibly because they act in similar biological pathways and thus if the pathway is already subverted by *BRCA1* or *BRCA2* mutations there is no additional advantage conferred by mutations in additional DNA repair genes. These results are different to the common, low-penetrance breast cancer SNP which do confer an increased risk in *BRCA2* carriers (*16*).

**Key Research Accomplishments**

In this reporting period we have achieved the following:

- We published a review entitled Breast Cancer Predisposition – Past, Present and Future in Annual Review of Genomics and Human Genetics, the leading review journal in this area. We were notified in Dec 2008 that our article was in the top 5 to be accessed from their website and that the abstract had been accessed 1127 times and the PDF has been downloaded 647 times.

- We have performed a Genome-wide association scan in 4000 familial breast cancer cases and 6000 controls.

**Reportable Outcomes**

1. Turnbull C and Rahman N (2008) Genetic predisposition to Breast cancer: past, present and future. *Annual Review of Genetics and Genomics* 9:321-45
2. Rahman N, invited speaker at American Society of Human Genetics session on Breast Cancer Genetics, Philadelphia, 2008.
3. Based in part on the work supported by this award we successfully applied to the Wellcome Trust to fund our second generation genome-wide association study - £655,500, which is currently underway.
4. Dr Muna Ahmed, a clinical research fellow supported on this programme was awarded her research doctoral degree based on the research she undertook as part of this award. She subsequently successfully applied for a Specialist clinical fellowship in Cancer Genetics.

## Conclusion

We have had another productive year. A considerable part of this reporting period was spent putting together the 4000 familial samples for our GWAS scan and we complete and replicate the findings from this in the next reporting period. We have continued our sequencing endeavors to clarify the role of *GEN1* in breast cancer susceptibility and to investigate the interactions of rare, intermediate-penetrance breast cancer susceptibility alleles with *BRCA1* and *BRCA2* mutations. This has revealed interesting differences to the common, low-penetrance alleles. In the next year we aim to begin exome-wide sequencing analyses using next generation sequencing technologies to continue to investigate the genetic basis of breast cancer.

**References**


1. M.R. Stratton, N. Rahman N, The emerging landscape of breast cancer susceptibility. *Nature Genetics* **40**, 17-22 (2008).

2. C. Turnbull, N. Rahman, Genetic predisposition to breast cancer, past, present, and future. *Annu.Rev Genomics Hum Genet* **9**, 321-45 (2008).

3. R. Wooster, G. Bignell, J. Lancaster, S. Swift, S. Seal, *et al*., Identification of the breast cancer susceptibility gene BRCA2 *Nature* **378**, 789-92 (1995).

4. H. Meijers-Heijboer, A. van den Ouweland, J. Klijn, M. Wasielewski, A. de Snoo, *et al.,* Low-penetrance susceptibility to breast cancer due to CHEK2*1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nature Genetics* **31**, 55-59 (2002).

5. S. Seal, D. Thompson, A. Renwick, A. Elliott, P. Kelly, *et al.* Truncating mutations in BRIP1 are low penetrance breast cancer susceptibility alleles. *Nature Genetics* **38**, 1239-1241 (2006).

6. A. Renwick, D. Thompson, S. Seal, P. Kelly, T. Chagtai, *et al.* ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. *Nature Genetics* **38**, 873-875 (2006).

7. N. Rahman, S. Seal, D. Thompson, P. Kelly, A. Renwick, *et al.* PALB2, which encodes a BRCA2 interacting protein, is a breast cancer susceptibility gene. *Nature Genetics* **39**, 165-167 (2007).

8. D. F. Easton, K. A. Polley, A. M. Dunning, P. D. Pharaoh, D. Thompson, *et al*. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* **447**, 1087-1093 (2007).

9. Wellcome Trust Case Control Consortium and The Australo-Angle-American Spondylitis Consortium Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nature Genetics* **39**, 1329-1338 (2007).

10. M. Garcia-Closas, S. Chanock, Genetic susceptibility loci for breast cancer by estrogen receptor status. *Clin Cancer Res* **14**,8000-9 (2008).

11. S. Ahmed, Thomas G, Ghoussaini M, Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2 *Nature Genetics* (in press, 2009)

12. D. J. Hunter, P. Kraft, K.B. Jacobs, D. G. Cox, M. Yeager, *et al*. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nature Genetics* **39**, 870-4 (2007).

13. S. N. Stacey, A. Manolescu, P. Sulem, T. Rafnar, J. Gudmundsson, *et al*. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nature Genetics* **39**, 865-9 (2007).

14. S. P. Ip, U. Rass, M. G. Blanco, H. R. Flynn, J. M. Skehel, *et al.* Identification of Holliday junction resolvases from humans and yeast. *Nature* **456**, 357-61 (2008).

15. L. D. Wood, D. W. Parsons, S. Jones, J. Lin, T. Sjoblom, R. J. Leary. The genomic landscapes of human breast and colorectal cancers. *Science* **318**, 1108-13 (2007).

16. A. C. Antoniou, A. B. Spurdel, O.M. Sinilnikova, S. Healey, K. A. Pooley, *et al*. Common breast cancer-predisposition alleles are associated with breast cancer risk in BRCA1 and BRCA2 mutation carriers. *Am J Hum Genet* **82**, 937-48 (2008).

## Appendices

Attached paper - Turnbull C and Rahman N (2008) Genetic predisposition to Breast cancer: past, present and future. *Annual Review of Genetics and Genomics* 9:321-45

## Supporting Data

None

# Genetic Predisposition to Breast Cancer: Past, Present, and Future

Clare Turnbull and Nazneen Rahman

Section of Cancer Genetics, Institute of Cancer Research, Sutton, SM2 5NG, United Kingdom; email: clare.turnbull@icr.ac.uk, nazneen.rahman@icr.ac.uk

Annu. Rev. Genomics Hum. Genet. 2008. 9:321–45

First published online as a Review in Advance on June 10, 2008

The *Annual Review of Genomics and Human Genetics* is online at genom.annualreviews.org

This article's doi: 10.1146/annurev.genom.9.081307.164339



1527-8204/08/0922-0321$20.00

## Key Words

BRCA*, allele, familial, susceptibility, penetrance

## Abstract

In recent years, our understanding of genetic predisposition to breast cancer has advanced significantly. Three classes of predisposition factors, categorized by their associated risks of breast cancer, are currently known. *BRCA1* and *BRCA2* are high-penetrance breast cancer predisposition genes identified by genome-wide linkage analysis and positional cloning. Mutational screening of genes functionally related to *BRCA1* and/or *BRCA2* has revealed four genes, *CHEK2*, *ATM*, *BRIP1*, and *PALB2*; mutations in these genes are rare and confer an intermediate risk of breast cancer. Association studies have further identified eight common variants associated with low-penetrance breast cancer predisposition. Despite these discoveries, most of the familial risk of breast cancer remains unexplained. In this review, we describe the known genetic predisposition factors, expound on the methods by which they were identified, and consider how further technological and intellectual advances may assist in identifying the remaining genetic factors underlying breast cancer susceptibility.

## INTRODUCTION

**Linkage analysis:** a statistical method for evaluating linkage between disease and markers of known location by following their inheritance in families

**Penetrance:** the probability that a particular phenotype/disease is expressed in an individual with a particular genotype

Through the dynamic interplay of multiple approaches, the past two decades have witnessed the gradual emergence of a clearer understanding of genetic susceptibility to breast cancer. Clinical observation underpinned these advances through recognition of unusual familial clustering and phenotypes associated with breast cancer. Observational epidemiology has also provided essential foundations to our understanding, both in quantifying the contribution of genetic factors to breast cancer and in predicting how these factors will act individually and interact. Genetic modeling has been performed to predict the profiles of the genes involved, which in turn facilitates the selection of molecular methods appropriate for identifying them. Linkage analysis, mutational screening of candidate genes, and association studies have been used to identify predisposition factors of three distinct risk-prevalence profiles: rare high-penetrance alleles, rare intermediate-penetrance alleles, and common low-penetrance alleles.

## OBSERVATIONAL EPIDEMIOLOGY AND SEGREGATION ANALYSIS

Epidemiological observation of the clustering of breast cancer within families was made in Roman times and reiterated by other early socio-scientific commentators (22, 77). More than 50 studies have explored this familial aggregation of breast cancer using predominantly case-control and cohort designs. Meta-analyses of these data conclude that, overall, breast cancer is twice as common in women with an affected first-degree relative. Simulation studies suggest that an environmental factor would have to confer at least a tenfold increase in risk for shared exposure to result in even a modest increase to the familial risk (65). Because no environmental risk factors of this magnitude have been identified for breast cancer, it seems unlikely that shared environment accounts for much of the familial aggregation. Heritability studies can assist in clarifying the relative contribution of genetic factors and shared environment. Via demonstration that the risk to a monozygotic twin is substantially higher than to a dizygotic twin of an affected individual, heritability studies using twins confirmed that the predominant component of the familial aggregation in breast cancer is genetic (80, 93). Patterns of disease aggregation in twins provide evidence not just for a genetic basis for breast cancer but demonstrate a markedly skewed distribution of genetic liability, suggesting that the majority of genetic risk may lie within a genetically predisposed minority (93).

Segregation analysis (genetic modeling) involves the simulation of various scenarios whereby disease occurs owing to combinations of known or hypothetical genes of specified risk, prevalence, and mode of inheritance. The patterns of cancers predicted by the different models are compared with observed disease frequencies and the best-fitting model is favored. Segregation analyses have strongly influenced the molecular approaches with which genetic predisposition factors have been investigated. The Cancer and Steroid Hormone Study data and other early segregation analyses of families with breast cancer favored a highly penetrant autosomal dominant genetic model (16, 29, 68, 121). This prediction was validated through linkage analysis and the identification of the highly penetrant autosomal dominant breast cancer predisposition genes, *BRCA1* and *BRCA2*. Subsequent complex segregation analyses have incorporated the effects of *BRCA1* and *BRCA2* and explored the evidence for additional genes (8, 9, 35, 89). Probably the most comprehensive of these analyses, based upon an extensive series comprising both population-based cases of breast cancer and large familial clusters, has strongly favored the polygenic model. The polygenic component of risk in this model is equivalent to multiple genetic factors acting independently: This component is log-normally distributed, has a variance that declines linearly with age, and applies similarly to *BRCA* mutation carriers and noncarriers (8–10). The polygenic model is also

consistent with multiple observations that the excess of familial breast cancer is distributed across many families, each typically comprising a modest number of cases, rather than just a few very extensive families (9, 35). The polygenic model has largely been accepted as an apposite explanation for the residual breast cancer predisposition and has been validated in part by identification of a number of the incumbent polygenes.

## STRATEGIES FOR IDENTIFYING BREAST CANCER PREDISPOSITION FACTORS

Three principal experimental designs have been used in the molecular identification of genetic breast cancer predisposition factors: genome-wide linkage analysis, mutational screening of candidate genes, and association studies.

### Linkage Analysis and Positional Cloning

Linkage studies are used to map a disease locus via analysis for cosegregation of genomic markers with a specified disease phenotype using samples from multiple members of large families. The region of the genome surrounding the linked markers is then interrogated for likely causative genes (positional cloning). Linkage analysis is suitable for mapping only high-penetrance breast cancer predisposition genes, whereby mutations result in prominent cancer families in which the majority of affected individuals carry the mutation. If a gene is of lower penetrance, the correlation between breast cancer and mutation status in mutation-positive families may be insufficient to generate a linkage signal. Nor will a significant signal be detectable if the breast cancer is linked to a particular locus in only a small proportion of the families analyzed.

Linkage maps and sets of markers that adequately cover the genome became available in the late 1980s. Using large breast cancer pedigrees, genome-wide linkage analyses were undertaken to map the high-penetrance breast cancer susceptibility genes proposed by the segregation analyses, *BRCA1* and *BRCA2* (61, 123). Linkage analysis has also been performed in groups of phenotypically distinct families to map the loci underlying specific syndromes that include an elevated risk of breast cancer. Positional cloning was then used to identify the causative genes, such as *PTEN*, *STK11*, and *CDH1* (58, 63, 64, 70, 87, 88).

### Resequencing Studies: Mutational Screening of Candidate Genes

During the 1990s, emerging understanding of the molecular pathogenesis of breast cancer offered insights into possible candidate genes that predispose to breast cancer. It seemed biologically plausible that proteins interacting with BRCA1 and BRCA2 or acting in similar DNA repair pathways may also be involved in breast cancer susceptibility. A few deleterious mutations were found at frequencies amenable to evaluation by association studies. However, in the majority of DNA repair genes studied, disease-causing mutations (primarily leading to premature protein truncation or nonsense-mediated RNA decay) have been individually very rare. A robust demonstration that such genes confer predisposition to breast cancer has therefore required mutational screening of the entire coding sequence of the gene through large numbers of cases and controls to meaningfully compare the total number of pathogenic mutations. Experiments of this magnitude have been published for only a handful of DNA repair genes. Other postulated candidate genes, such as those implicated in cell cycle regulation, checkpoint control, apoptosis, and steroid hormone metabolism, have rarely been evaluated to this level.

### Association Studies

In an association study, the frequency of a specified variant is compared between breast cancer cases and controls. Statistically significant differences in allele frequency between cases and controls are more readily demonstrable for

**Single nucleotide polymorphism (SNP):** a single-base sequence variation that commonly occurs at a particular position within the genome

**Linkage disequilibrium:** the non-random association of alleles at two or more loci, which suggests that they may be physically close and thus linked

variants of population frequency >5%; hence association studies are most suitable for identification of common breast cancer predisposition variants. Initial studies focused on genes proposed by function and examined the association with breast cancer of numerous common variants from recognized and candidate predisposition genes. However, experiments were often too small, subject to bias, and utilized too lenient levels of significance, resulting in inconsistency and lack of replication of findings. The majority of putative associations likely represented false positives (type 1 errors) (40, 95). To increase power and rationalize choices of candidates, collaborations have formed that undertake association studies across tens of thousands of samples (19).

The emergence of comprehensive high-density maps of single nucleotide polymorphisms (SNPs) and affordable genotyping platforms has allowed the graduation of association studies from the limitations of preconceived notions of candidacy to the agnosticism of the genome-wide approach. On account of linkage disequilibrium, a panel of a few hundred thousand reporter SNPs can be used as tags for the majority of the millions of common variants in the genome. Accordingly, owing to the degree of multiple testing, a genome-wide scan must be sufficiently well-powered to ensure that the true associations are detected. The staged experimental design is a means of optimizing the statistical power afforded by a given sample size and has been useful in studies to date because of the current high costs of whole genome tag SNP panels.

## BREAST CANCER PREDISPOSITION GENES AND VARIANTS

The breast cancer predisposition factors identified to date can be stratified by risk profile into three tiers: high-penetrance genes, intermediate-penetrance genes, and low-penetrance alleles. Three further genes are associated with syndromes in which the incidence of breast cancer is elevated but the actual risk remains unclear (**Table 1**).

### High-Penetrance Breast Cancer Predisposition Genes

Mutations in three high-penetrance breast cancer predisposition genes confer a greater than tenfold relative risk of breast cancer. *BRCA1* and *BRCA2* were identified through linkage analysis and positional cloning. *TP53* was deemed a plausible candidate and identified as a high-risk breast cancer gene through mutational screening.

**BRCA1 and BRCA2.** The first convincing report of linkage of breast cancer to 17q21 was published in 1990 (61). In 1994, positional cloning revealed the causative gene, accordingly named *BRCA1* (86). Linkage analysis and positional cloning led to the mapping and identification of *BRCA2* in 1994 and 1995, respectively (122, 123). BRCA1 and BRCA2 have important roles in the maintenance of genomic stability by facilitating repair of DNA double-strand breaks. The cellular roles of BRCA1 and BRCA2 were reviewed by Gudmundsdottir & Ashworth (56).

*BRCA1* and *BRCA2* are large genes in which multiple different loss-of-function mutations have been detected. Some founder mutations are relatively frequent in particular ethnic groups, such as *BRCA1*_185delAG, *BRCA1*_5382insC, and *BRCA2*_6174delT in the Ashkenazim and *BRCA2*_999del5 in Icelanders. However, the majority of mutations are individually rare and many have been reported only in single families. Most recognized disease-associated mutations result in premature protein truncation and include nonsense mutations, deletions/insertions that result in translational frameshifts, and mutations that affect splice sites. More recently a number of exonic deletions/duplications have also been identified, especially in *BRCA1*. In addition, a large number of amino acid substitutions and synonymous nucleotide substitutions in *BRCA1* and *BRCA2* have been

**Table 1 Summary of known breast cancer predisposition factors**

| | Gene/Locus | Relative Risk of breast cancer | Carrier Frequency† | Breast cancer subtype | Other cancers in monoallelic carriers | Syndrome in biallelic carriers | Method of identification |
|---|---|---|---|---|---|---|---|
| High penetrance | *BRCA1* | >10 | 0.1% | Basal (ER-negative) | Ovarian | | Linkage study |
| | *BRCA2* | >10 | 0.1% | | Ovarian prostate | Fanconi anaemia D1 | Linkage study |
| | *TP53* | >10 | rare | | Sarcomas adrenal brain | | Candidate resequencing study |
| Uncertain penetrance | *PTEN* | 2–10 | rare | | Thyroid endometrium | | Linkage study |
| | *STK11* | 2–10 | rare | | Gasto-intestinal | | Linkage study |
| | *CDH1* | 2–10 | rare | lobular | Gastric (diffuse) | | Linkage study |
| Intermediate penetrance | *ATM* | 2–3 | 0.4% | | | Ataxia telangiectasia | Epidemiology; Candidate resequencing study |
| | *CHEK2* | 2–3 | 0.4% | | | | Candidate resequencing study |
| | *BRIP1* | 2–3 | 0.1% | | | Fanconi anaemia J | Candidate resequencing study |
| | *PALB2* | 2–4 | rare | | | Fanconi anaemia N | Candidate resequencing study |
| Low penetrance | 10q26, 16q12, 2q35, 8q24, 5p12 | 1.08–1.26 | 24–50% | ER-positive | | | Genome-wide association studies |
| | 11p15, 5q11 | 1.07–1.13 | 28–30% | | | | Genome-wide association study |
| | 2q33 | 1.13 | 0.87 | | | | Candidate association study |

†estimated carrier frequency of mutations/risk allele in the UK; where 'rare', the carrier frequency is unlikely to be >0.1%.

detected, the majority of which are innocuous. A few missense mutations, for example those that target cysteine residues in the BRCA1 RING domain, abrogate function and are presumed to confer risks comparable to truncating mutations (43, 106).

There is evidence of genotype-phenotype correlations for mutations in *BRCA1* and *BRCA2*. In both genes, the ratio of ovarian cancers to breast cancers conferred by mutations in the central region of the gene is higher than that for mutations in the 5′ or 3′ end. In *BRCA1* this central region is bounded by nucleotides 2401 and 4191; in *BRCA2* the ovarian cancer cluster region lies between nucleotides 3035 and 6629 (114, 115). The biological mechanisms underlying these observations remain opaque (120).

*BRCA1* and *BRCA2* are high penetrance breast cancer genes. Estimates of the risks of cancer conferred by mutations in these genes vary according to the ascertainment of the cases studied. Early studies of large cancer families suggested that the risk of breast cancer by age 70 may be as high as 87% [95% confidence interval (CI) = 72%–95%] for *BRCA1* and 84% (43%–95%) for *BRCA2* mutation carriers, although the upper estimates were based on relatively small numbers of families (44, 51). In population-based studies of breast cancer cases, unselected for family history, the risks are lower: 65% (51%–75%) for *BRCA1* and 45% (33%–54%) for *BRCA2* (5). The pattern of age-related risk in *BRCA2* mutation carriers resembles that of the general population (only higher). By contrast, the relative risk of breast cancer is markedly elevated in *BRCA1* mutation carriers under 40 and becomes less dramatic with advancing age (5). Linkage analysis suggests that mutations in *BRCA1* and *BRCA2* are responsible for disease in approximately two-thirds of large families with site-specific female breast cancer ($\geq$ four cases) but the attribution diminishes sharply for smaller family clusters (51). The estimated population frequency of mutations in these genes is approximately 1/1000 per gene in the United Kingdom. Overall this equates to 15%–20% of the excess familial risk of breast cancer (4, 7, 9, 42, 91).

*BRCA1* and *BRCA2* are also high-penetrance ovarian cancer genes: Mutations in *BRCA1* confer a higher risk of ovarian cancer than those in *BRCA2*, particularly for carriers below 50 years of age (5, 7, 44, 51). *BRCA1* and *BRCA2* mutations account for most of the epidemiologically observed familial coaggregation of breast and ovarian cancer (45, 92). 95% of families containing four or more breast cancers and an ovarian cancer were linked to *BRCA1* or *BRCA2*, whereas recent mutational screening detected mutations in 83% of families containing at least two cases of each cancer (51, 99).

The relative risk of male breast cancer is elevated for both genes, particularly *BRCA2*. An elevated risk of prostate cancer has also been demonstrated in *BRCA2* carriers, particularly in men aged <65 years (21, 116). Small excesses of a number of other cancers have been observed in monoallelic (heterozygous) *BRCA1* and *BRCA2* mutation carriers but larger studies are required to clarify whether these findings reflect truly elevated risks of these cancers (21, 116). Biallelic mutations in *BRCA2* result in Fanconi anemia, subtype D1. Fanconi anemia is a rare, recessive chromosomal instability syndrome characterized by skeletal abnormalities, bone marrow failure, and cancer predisposition. Subtype D1 has a distinctive phenotype that includes a high risk of childhood solid tumors (66). Biallelic *BRCA1* mutations have never been reported convincingly in humans and are presumed to be embryonically lethal (49).

*BRCA1* tumors are typically high grade, invasive ductal carcinomas in which there is a high incidence of triple negative phenotype: negative staining for ER (estrogen receptor), PR (progesterone receptor), and HER2 (ERBB2) (20, 73, 75). These tumors also frequently stain positively for a subset of basal keratins characteristically expressed in the normal basal myoepithelium of the breast. This basal phenotype is distinctive and largely encompasses the previous histological descriptions of *BRCA1*

tumors as medullary (69, 74). No distinctive histopathological features have been described in *BRCA2* tumors.

**TP53.** Li-Fraumeni syndrome is a cancer predisposition syndrome in which there is a high frequency of early onset breast cancer found in association with sarcomas and childhood cancers of the adrenal cortex, brain, and other sites. Although of a penetrance sufficient for mapping through linkage analysis, the associated early mortality and rarity of the condition impeded collection of sufficient familial samples. p53 was recognized early as a prominent transcription factor central to multiple cellular pathways and is frequently somatically mutated in tumors. These observations recommended *TP53* as a plausible candidate gene for Li-Fraumeni syndrome and in 1990 mutational screening of the gene revealed causative mutations in the five families studied (82). The overall lifetime cancer risk for women with Li-Fraumeni syndrome is grossly elevated, predominantly on account of their high risk of breast cancer (15, 28). Li-Fraumeni syndrome is rare and mutations in *TP53* are uncommon in non-Li-Fraumeni breast cancer families (18, 48, 76). Thus, the attributable risk of *TP53* mutations to familial breast cancer is very low.

## Breast Cancer Predisposition Genes of Uncertain Penetrance

Three syndromes are currently clearly associated with an increased risk of breast cancer, but the magnitude of the associated risk for each remains uncertain: Cowden syndrome (caused by *PTEN* mutations), Peutz-Jeghers syndrome (caused by *STK11* mutations), and Hereditary diffuse gastric cancer syndrome (caused by *CDH1* mutations).

**PTEN.** Cowden syndrome is a multiple hamartoma syndrome that includes increased risk of benign and malignant tumors of the breast, thyroid, and endometrium; distinctive higher-penetrance features include mucocutaneous lesions, macrocephaly, and hamartomatous intestinal polyps. A study of 12 Cowden syndrome families revealed linkage to chromosome 10q, leading to identification of *PTEN* as the causative gene (87, 88). *PTEN* encodes a lipid phosphatase that functions as a tumor suppressor through negative regulation of a cell-survival signaling pathway. There is cross talk between this PTEN-related pathway and other pathways, including those involving Ras, p53, and TOR (36).

**STK11.** Peutz-Jeghers syndrome is characterized by hamartomatous intestinal polyps, mucocutaneous pigmentation, and increased incidence of several malignancies, including breast cancer. Directed by patterns of loss of heterozygosity in polyps of affected individuals, studies in 12 Peutz-Jeghers families established linkage to chromosome 19p and positional cloning led to the identification of *STK11* (*LKB1*) as the causative gene. STK11 is a serine/threonine kinase that inhibits cellular proliferation, controls cell polarity, and interacts with the TOR pathway (1, 63, 64, 70).

**CDH1.** Linkage analysis in a single New Zealand Maori family with multiple cases of diffuse gastric cancer enabled the identification of *CDH1* (*ECAD*) as the responsible gene (58). Many additional families have since been reported and an elevated frequency of lobular breast carcinoma has been observed (71, 96). Occasional *CDH1* mutations in families with cases of lobular breast cancer but no gastric cancer have also been reported (83). *CDH1* encodes E-cadherin, a transmembrane protein important in the maintenance of cell polarity.

These conditions were identified because the non-breast-cancer-related features were sufficiently distinctive and penetrant to reliably ascertain families and assign affection status, thus facilitating mapping of the causative genes by linkage analysis. The true breast cancer risks associated with mutations are unclear and it is possible that the published risks are inflated through the bias of studying families with prominent phenotypes. However, the relative risks are likely to be intermediate and in the

range of 2 to 10. There is no current evidence that mutations in *PTEN*, *STK11*, or *CDH1* account for a substantial proportion of familial or sporadic breast cancer in the absence of their respective syndromes; because these syndromes are very rare, the attributable risk of mutations in these genes to familial breast cancer is low (14, 27, 57, 83, 107).

## Intermediate-Penetrance Breast Cancer Predisposition Genes

Four intermediate-penetrance breast cancer genes have been identified via mutational screening: *CHEK2*, *ATM*, *BRIP1*, and *PALB2*. Mutations in these genes are rare and confer a relative risk of breast cancer of 2 to 4. *RAD50* may also be an intermediate-penetrance breast cancer predisposition gene, but convincing results to date pertain only to a founder mutation detected in the Finnish population. The rarity of mutations and modest associated risks are such that the attributable risk of mutations in these genes is low: Together they account for approximately 2.3% of excess familial risk (98).

**CHEK2.** *CHEK2* encodes CHK2, a central mediator of cellular response to DNA damage that phosphorylates both p53 and BRCA1 to regulate repair of DNA double-strand breaks. The 1100delC mutation in *CHEK2* was first reported after mutational screening of this eligible candidate gene in a single family with features of Li-Fraumeni syndrome (13). However, the population frequency of this mutation was shown to be approximately 1% (18/1620), demonstrating that *CHEK2* could not be a high-risk Li-Fraumeni syndrome gene. The identification of the *CHEK2*_1100delC mutation in 4.2% (30/718) of breast cancer families demonstrated, instead, that *CHEK2* was an intermediate-penetrance breast cancer predisposition gene ($P = 5 \times 10^{-6}$) (85). Segregation analysis in these families provided an indirect estimate of 1.70 (95% CI = 1.32–2.20) for the relative risk of breast cancer conferred by the mutation. Combining data from ten case control studies, the *CHEK2* consortium demonstrated the frequency of *CHEK2*_1100delC in population-based breast cancer cases to be 1.9% (201/10,860) compared with 0.7% in controls (64/9065) ($P = 1 \times 10^{-7}$). These frequencies equate to a direct odds ratio of breast cancer of 2.34 (95% CI = 1.72–3.20) (26). It has been proposed that *CHEK2*_1100delC may confer risks of prostate and other cancers but evidence is conflicting (37, 38, 85, 105, 117). There are also reports of other rare truncating mutations in *CHEK2*, which would be anticipated to have similar risks to *CHEK2*_1100delC (17, 39, 55, 103, 119). A number of rare *CHEK2* missense variants, such as 470T/C (I157T), have been reported in breast cancer cases but their significance has yet to be clarified (2, 17, 55, 72, 103).

**ATM.** The proposal of *ATM* as a breast cancer predisposition gene first came in 1976 from an epidemiological study that reported an excess of breast cancer in female relatives of patients with ataxia telangiectasia, an autosomal recessive syndrome characterized by progressive cerebellar ataxia, immune deficiency, and cancer predisposition. This astute observation preceded the mapping of the gene by almost two decades and has been subsequently replicated in a number of large epidemiological studies (41, 102, 112, 113). The candidacy of *ATM* as a breast cancer predisposition gene was further enhanced as the function of the encoded protein became apparent: ATM occupies a central role in the response to double-strand DNA breaks through initiation of a signaling cascade that involves phosphorylation of multiple proteins including p53, BRCA1, and CHK2. Results from initial mutational screening of *ATM* in breast cancer were inconclusive and/or inconsistent but convincing proof was finally published in 2006. Mutations were found in 12/443 familial cases negative for mutations in *BRCA1* and *BRCA2* and 2/521 controls ($P = 0.0047$). These mutations comprised truncations, splice-site abnormalities, and two missense mutations that were known to affect protein function and

cause ataxia telangiectasia. The relative risk of breast cancer conferred by *ATM* mutations was estimated from segregation analysis to be 2.37 (95% CI = 1.51–3.78, $P = 0.0003$) (101). This indirect estimate is very similar to that derived from large epidemiological studies (113). There is some epidemiological evidence that mutations in *ATM* may predispose to other cancers but this has yet to be confirmed or corroborated by molecular data (113).

**BRIP1.** *BRIP1* (*BACH1*) encodes a DEAH helicase that interacts with BRCA1 and has BRCA1-dependent roles in DNA repair and checkpoint control (24, 90). In 2006, truncating *BRIP1* mutations were reported in 9/1212 index breast cancer cases from families negative for mutations in *BRCA1* and *BRCA2* compared with 2/2081 controls, thus providing convincing evidence that *BRIP1* is a breast cancer predisposition gene ($P = 0.0030$). Segregation analysis found the relative risk conferred by *BRIP1* mutations to be 2.0 (95% CI = 1.2–3.2, $P = 0.012$) (104). Of the mutations reported in *BRIP1*, approximately half are a nonsense mutation, 2392C/T (R798X), whereas the remainder include disparate small insertions and deletions. Concurrently, it emerged that biallelic mutations in *BRIP1* result in Fanconi anemia, although the Fanconi anemia subtype associated with *BRIP1* (subtype J) is phenotypically distinct from that associated with *BRCA2* (subtype D1) and has not been associated with childhood solid tumors (78, 79, 81).

**PALB2.** Precipitation of BRCA2-containing complexes revealed a novel protein that was shown to promote the localization and stability of BRCA2, thus facilitating BRCA2-mediated DNA repair. Further study of *PALB2* (partner and localizer of BRCA2) revealed that knockdown of the gene resulted in sensitization of cells to chromosomal damage by mitomycin C, the hallmark of Fanconi anemia (125). Biallelic truncating mutations in *PALB2* were detected in families with Fanconi anemia not caused by mutations in other genes and this subtype was designated FA-N (100, 124). The phenotypes of FA-N and Fanconi anemia caused by *BRCA2* mutations (FA-D1) are very similar and can be distinguished from classical Fanconi anemia on account of a markedly elevated frequency of childhood solid tumors such as Wilms tumor and medulloblastoma. With the recent precedents of both *BRIP1* and *BRCA2* causing Fanconi anemia in biallelic mutation carriers and conferring susceptibility to breast cancer in monoallelic carriers, the identification of *PALB2* as a Fanconi anemia gene further recommended *PALB2* as an attractive candidate breast cancer predisposition gene.

Sequencing of the gene revealed truncating mutations in 10/923 index breast cancer cases from families negative for mutations in *BRCA1* and *BRCA2* compared with 0/1084 in controls ($P = 0.0004$). Segregation analysis from these families estimated the relative risk of *PALB2* mutations to be 2.3 (95% CI = 1.4–3.9, $P = 0.0025$) (98). Mutational screening of *PALB2* in a Finnish series detected one truncating mutation: 1592delT. This mutation was identified in 3/113 (2.7%) familial cases, 18/1918 (0.9%) unselected breast cancer cases, and 6/2501 (0.2%) controls. The directly derived odds ratio of breast cancer of this mutation is therefore 3.94 (95% CI = 1.5–12.1) (47). This mutation has not been detected in other series and may represent a Finnish founder. A French Canadian founder mutation, 2323 C/T (Q775X), has also been reported (52).

**RAD50.** The highly conserved MRN complex (*MRE11*, *RAD50*, and *NBS1*) interacts with BRCA1 and plays a central role in DNA repair. A truncating mutation in *RAD50*, 657delT, was demonstrated in 8/317 consecutively ascertained breast cancer cases and 6/1000 controls from Finland [$P = 0.008$, odds ratio (OR) = 4.3, 95% CI = 1.5–12.5]. Other rare truncating mutations have been reported in *RAD50*, but the contribution of *RAD50* to breast cancer predisposition outside of Finland requires further clarification (62, 118).

**Distinctive features of intermediate-penetrance breast cancer predisposition genes.** As in *BRCA1* and *BRCA2*, multiple different truncating mutations occur in these intermediate-penetrance genes, the frequencies of which differ between populations on account of founder and migrational effects. Monoallelic mutations in these genes confer an approximately twofold increase in the risk of breast cancer. Although there is some imprecision in the indirect risk estimates derived from segregation analysis, population-based series and epidemiological data have provided some corroboration for these figures. There is currently no strong evidence that monoallelic mutations in these genes confer a phenotype beyond breast cancer predisposition. Biallelic mutations in *ATM*, *BRIP1*, and *PALB2* result in severe childhood disorders: ataxia telangiectasia and Fanconi anemia types J and N, respectively (97). Because these genes function in the same pathways as BRCA1 and BRCA2, it remains unclear why abrogation of their functions confers more modest risks of breast cancer. However, the lower penetrance of these genes means that they deviate from some of the hallmarks of pathogenicity associated with mutations of *BRCA1* and *BRCA2*. Mutations in these genes do not show the classical, near-complete pattern of segregation of disease with the family mutation. By contrast, because breast cancer is a common disease, in a typical pedigree an appreciable proportion of breast cancers arise by chance and thus may manifest in either mutation carriers or noncarriers. Likewise, one should not anticipate finding the pattern of loss of the wild-type allele detected at high frequency in tumors in *BRCA1* and *BRCA2* mutation-positive individuals. Firstly, the mechanism of tumorigenesis of these genes is unknown and may not require inactivation of the wild-type allele. Secondly, even if the classical, two-hit model of a tumor suppressor gene does apply, one could not expect to demonstrate loss of heterozygosity in all tumors occurring in mutation carriers because the mutation is not causally implicated in a significant portion of them.

## Low-Penetrance Breast Cancer Predisposition Alleles

There is currently strong evidence for the association with breast cancer of eight common alleles, which each confer a relative risk of breast cancer of <1.5. A nonsynonymous variant in *CASP8* was identified through a candidate-based approach in a consortium comprising 14 studies that compared 17,109 cases and 16,423 controls (34). Seven further variants have been detected in three recent genome-wide association studies (**Table 2**). Easton and coworkers (46) performed a three-stage genome-wide experiment. In the first stage a Perlegen platform was used to study the association of 227,876 SNPs in 390 cases with a family history of breast cancer and/or bilateral disease and 364 controls from the United Kingdom. In the second stage 12,711 SNPs (the most significant 5% of SNPs from stage 1) were genotyped in 3990 cases and 3916 UK controls using a custom-designed array. In the third stage the 30 most significant SNPs were tested for confirmation in 21,860 cases and 22,578 controls from 22 international studies (46). The study identified five variants associated with breast cancer at a significance level of $P < 10^{-7}$. Stacey and colleagues (110) conducted a genome-wide scan in 1600 Icelandic cases and 11,563 controls using the Illumina HumanHap300 platform; the ten best associated SNPs were studied further in five international replication sets comprising 2954 cases and 6014 controls. This study identified one of the variants reported by Easton and coworkers (rs3803662) and a novel region of association on 2q35. As part of the National Cancer Institute Cancer Genetic Markers of Susceptibility (CGEMS) Project, Hunter and colleagues (67) used the Illumina HumanHap500 array to genotype 1145 postmenopausal invasive breast cancer cases and 1142 North American controls of European ancestry. The six best-associated signals were

**Table 2 Summary of known low-penetrance breast cancer predisposition variants***

| Locus | Estimated size of LD block (kb) | Genes within LD block | SNP | MAF | Heterozygote OR (95% CI) | Homozygote OR (95% CI) | Per allele OR (95% CI) | *P*- trend | Ascertainment | Study |
|---|---|---|---|---|---|---|---|---|---|---|
| 10q26 | 25 | *FGFR2* | rs2981582 | 0.38 | 1.23 (1.18–1.28) | 1.63 (1.53–1.72) | 1.26 (1.23–1.30) | $2 \times 10^{-76}$ | GW | Easton et al. (UK) (46) |
| | | | rs1219648 | 0.39 | 1.20 (1.07–1.42) | 1.64 (1.42–1.90) | | $1.1 \times 10^{-10}$ | GW | Hunter et al. (USA) (67) |
| 16q12 | 160 | *TNRC9* *LOC643714* | rs3803662 | 0.25 | 1.23 (1.18–1.29) | 1.39 (1.26–1.45) | 1.20 (1.16–1.24) | $1 \times 10^{-36}$ | GW | Easton et al. (UK) |
| | | | rs3803662 | 0.27 | 1.27 (1.19–1.36) | 1.64 (1.45–1.85) | 1.28 (1.21–1.35) | $5.9 \times 10^{-19}$ | GW | Stacey et al. (Iceland) (110) |
| 2q35 | 117 | – | rs13387042 | 0.50 | 1.11 (1.03–1.20) | 1.44 (1.30–1.58) | 1.20 (1.14–1.26) | $1.3 \times 10^{-13}$ | GW | Stacey et al. (Iceland) (110) |
| 5p12 | 310 | *MRPS30* | rs10941679 | 0.24 | | | 1.19 (1.13–1.26) | $2.9 \times 10^{-11}$ | GW | Stacey et al. (Iceland) (111); Easton et al. (UK) (46) |
| 5q11 | 280 | *MAP3K1* *MGC33648* *MIER3* | rs889312 | 0.28 | 1.13 (1.09–1.18) | 1.27 (1.19–1.36) | 1.13 (1.10–1.16) | $7 \times 10^{-20}$ | GW | Easton et al. (UK) (46) |
| 2q33 | 290 | *CASP8* *TRAK2* *ALS2CR12* *ALS2CR2* *ALS2CR11* *LOC389286* *LOC729191* | rs1045485 | 0.13 | 0.89 (0.85–0.94) | 0.74 (0.62–0.87) | 0.88 (0.84–0.92) | $1.1 \times 10^{-7}$ | Candidate | Cox et al. (UK) (34) |
| 8q24 | 110 | – | rs13281615 | 0.40 | 1.06 (1.01–1.11) | 1.18 (1.1–1.25) | 1.08 (1.05–1.11) | $5 \times 10^{-12}$ | GW | Easton et al. (UK) (46) |
| 11p15 | 180 | *LSP1* *TNNT3* *MRPL23* *H19* *LOC728008* | rs3817198 | 0.30 | 1.06 (1.02–1.11) | 1.17 (1.08–1.25) | 1.07 (1.04–1.11) | $3 \times 10^{-9}$ | GW | Easton et al. (UK) (46) |

*Abbreviations used: LD, linkage disequilibrium; MAF, minor allele frequency in control population; OR, odds ratio (measured relative to common homozygotes); CI, confidence interval; *P*-trend, *P* value for per allele effect under multiplicative model; GW, genome-wide association study; Candidate, association studies of single nucleotide polymorphisms (SNPs) in candidate genes.

further studied in 1776 cases and 2072 controls from three North American replication sets. The results of this study confirmed the signal in intron 2 of *FGFR2* reported by Easton and coworkers (67). The final predisposition SNP on 5p12 was identified through mining and cross-referencing of associations of borderline significance from these three genome-wide association studies; this signal was verified by Stacey and coworkers (111) in a replication series of 5,028 cases and 32,090 controls.

**10q26 (FGFR2 intron 2).** The strongest evidence for association was for a signal within intron 2 of the gene encoding fibroblast growth factor receptor 2 (*FGFR2*) [per allele OR = 1.26 (1.23–1.30) $P = 2 \times 10^{-76}$] (46, 67, 111). Fine association mapping of the region of linkage disequilibrium has been undertaken using UK samples and refined using an Asian series (in which linkage disequilibrium is weaker). These studies have reduced the association to a minimum set of six variants within intron 2, which are too strongly correlated for their individual effects to be discerned using further genetic epidemiologic approaches (46).

Somatic *FGFR2* mutations have been reported in several cancers and result in overactivity of the protein. It therefore seems plausible that the elevated breast cancer risk is somehow mediated through increased/altered activity of FGFR2, but the molecular basis for the association is currently unknown.

The risk-prevalence profile of this allele means that the power to detect it in the genome-wide scans is high. The corollary of this is that if other alleles of a similar risk-prevalence profile exist, they should most likely have been detected. Thus, it is possible that this may be the allele of most substantial effect to exist within this category of risk alleles. The differences in the risks conferred are nevertheless modest: The risk of breast cancer by age 70 in individuals homozygous for the risk allele is 10.5% compared with 6.7% for heterozygotes and 5.5% for nonrisk allele homozygotes (based upon UK breast cancer incidence figures). However, the risk allele is common: 14% of the UK population and 19% of UK breast cancer cases are homozygous for the risk allele. Hence, this allele may account for approximately 1.9% of the excess familial risk of breast cancer.

**16q12.** Strong evidence for association of rs3803662 with breast cancer was reported in two studies [per allele OR = 1.20 (1.16–1.24), $P = 1 \times 10^{-36}$ in the UK study and per allele OR = 1.28 (1.21–1.35), $P = 5.9 \times 10^{-19}$ in the Icelandic study] (46, 110). This variant lies on 16q12 and tags a region of linkage disequilibrium containing the 5′ end of *TNRC9* (*TOX3*) and a hypothetical gene *LOC643714*. *TNRC9* is a gene of uncertain function that contains a putative high mobility group box motif, suggesting that it may act as a transcription factor. Results from a previous study suggested that *TNRC9* expression is predictive of metastasis of breast cancer to bone (108).

**2q35.** The Icelandic study found evidence for association with breast cancer of a variant on 2q35 [per allele OR = 1.20 (1.14–1.26), $P = 1.3 \times 10^{-13}$]. The region of linkage disequilibrium does not contain any known genes. The nearest genes are *TNP1* (181 kb), *IGFBP5* (345 kb), and *IGFBP2* (376 kb) upstream and *TNS1* (761 kb) downstream. Although this signal was not detected in the other genome-wide scans, we have replicated the association in a UK series of familial breast cancer cases [per allele OR = 1.17 (95% CI 1.07–1.27), $P = 0.0004$] (C. Turnbull and N. Rahman, unpublished data) and it is being evaluated by the Breast Cancer Association Consortium.

**5p12.** Stacey and coworkers (111) noted that one of the ten top-ranked SNPs in the Icelandic genome-wide study was on chromosome 5p12 (rs7703618). The CGEMS study had reported an association of borderline significance on 5p12 at rs4866929, a SNP in tight linkage disequilibrium with rs7703618, while Easton and coworkers (46) had reported a tentative signal at 5p45 (rs981782), 371 kb away, which was also in loose linkage disequilibrium with rs7703618 ($r^2 = 0.10$). Struck by the

coincidence of tentative signals, they examined 21 SNPs in this region for association with breast cancer and found rs10941679 to be the most significantly associated SNP [per allele OR 1.19 (1.13–1.26), P = $2.9 \times 10^{-11}$]. The only gene in the region of linkage disequilibrium is *MRPS30* (*PDCD9*, programmed cell death protein 9), which encodes a component of the small subunit of the mitochondrial ribosome and has been implicated in apoptosis.

**5q11.** Association has been found for rs889312 [per allele OR = 1.13 (1.10–1.16), $P = 7 \times 10^{-20}$]. This SNP tags a 280-kb block of linkage disequilibrium on 5q11 that contains the genes *MAP3K1* (*MEKK*), *MGC33648*, and *MIER3*. *MAP3K1* is the most plausible candidate gene therein and encodes mitogen-activated protein kinase kinase kinase 1, which is involved in cell signaling (46).

**2q33 (CASP8 D302H).** This association was ascertained through the biological candidacy of *CASP8*, which encodes caspase-8, a protein involved in apoptosis [per allele OR = 0.88 (0.84–0.92), $P = 1.1 \times 10^{-7}$] (34). However, the region of linkage disequilibrium tagged by this SNP is approximately 290 kb in length and includes a number of other genes: *TRAK2* (encoding trafficking protein kinesin binding 2), three genes identified as candidates for juvenile amyotrophic lateral sclerosis 2 (*ALS2CR12*, *ALS2CR2*, and *ALS2CR11*), and two hypothetical genes (*LOC389286* and *LOC729191*). It is currently unclear whether D302H is the causal variant or whether it tags a distinct causal variant, which may or may not mediate its effect through CASP8. The relatively low minor allele frequency and risk of this allele mean that the power to detect its signal at the genome-wide level is comparatively low. This is consistent with failure of the allele to reach the arbitrary significance thresholds required for further investigation in the genome-wide association scans performed to date. It is also consistent with the presumed existence of many further alleles of comparable risk-prevalence profiles.

**8q24.** Association was found for a SNP within an 110-kb block of linkage disequilibrium on 8q24, which contains no known genes [per allele OR = 1.08 (1.05–1.11), $P = 5 \times 10^{-12}$] (46). It is interesting that in the first wave of genome-wide association studies in common cancers, 8q24 has also yielded multiple independent prostate cancer loci, one of which is also a colon cancer risk allele (rs6983267) (59, 60). The tag SNP associated with breast cancer does not demonstrate association with colon cancer or prostate cancer and is 60 kb proximal to rs6983267. Clustering of these predisposition alleles may be a coincidence or may indicate a common or related mechanism of cancer predisposition. The nearest gene to the breast cancer predisposition locus is *MYC* and it is plausible that the susceptibility occurs through some unknown mechanism of activation of this oncogene.

**11p15.** Another associated tag SNP lies in intron 10 of *LSP1* (*WP43*), which encodes lymphocyte-specific protein 1 [per allele OR = 1.07 (1.04–1.11), $P = 3 \times 10^{-9}$], an F-actin bundling cytoskeleton protein that is expressed in hematopoetic and endothelial cells (46). Other genes within this region of linkage disequilibrium include *TNNT3*, troponin T type 3; *MRPL23*, mitochondrial ribosomal protein; *H19*, an imprinted, maternally expressed, untranslated mRNA; and *LOC728008*, a hypothetical gene.

**Distinctive features of low-penetrance breast cancer predisposition alleles.** A fascinating aspect of the recent genome-wide scans has been the opaqueness of the relationship of the identified variants with known protein-coding genes. The patterns of linkage disequilibrium suggest that the causal variants need not lie in the coding region of a gene, as evidenced by the predisposition variants at 2q and 8q that are tens of kilobases away from the nearest protein-encoding gene. Any significance to the roles of genes upstream or downstream of the associated blocks of linkage disequilibrium is currently speculative.

The interpretation of subgroup analyses of the immunohistopathologic phenotypic variation of risk alleles requires caution on account of the limited power of some analyses. Initial analyses by the Icelandic group (110, 111), on identification of the predisposition SNPs at 16q12, 2q35, and 5p12, showed the effects of these alleles to be confined to ER-positive tumors. Although subsequent analyses of the SNPs at 10q26, 16q12, 8q24, 5q11, and 11p15 from the United Kingdom genome-wide association data also revealed that for all five alleles, the estimates of effect were stronger in ER-positive than ER-negative disease, the difference was only significant for those at 10q26 and 8q24. These two SNPs were also independently associated with a lower grade of disease. Adjusted analyses of these five SNPs did not reveal significant association with lymph node status nor survival (53).

Because these alleles occur at a high frequency in the general population, their population attributable risks (etiologic fractions) are relatively high (13%–16% for the alleles of stronger effects) (67, 110). However, this figure represents only the proportion of breast cancer cases in which the variant has played some causal role in development of disease. The associated risks are low and it is estimated that the five loci characterized by Easton and coworkers (46) account for a modest 3.6% of the excess familial risk of breast cancer in European populations.

## INTERACTIONS BETWEEN BREAST CANCER PREDISPOSITION FACTORS

Interaction is an important aspect of risk calculation, particularly in familial disease clusters in which multiple predisposition factors are likely active. The combined risk of two factors is strongly dependent upon the nature of the interaction between them. Typically the default model assumes that they act independently and multiplicatively (as per terms in a multiple regression analysis). The true situation may be more complex and comprise a mixture of multiplicative, additive, antagonistic, synergistic, and/or complex intermediate interactions.

The study of co-occurrence of genetic predisposition factors and exposition of gene-gene interactions has become more viable since the identification of convincing common predisposition alleles. Preliminary analyses from genome-wide association studies suggest that the low-penetrance risk alleles act multiplicatively with each other (46, 110). However, association studies of the SNPs at 10q26 and 5q11 performed within *BRCA1* and *BRCA2* mutation-positive families suggest that these SNPs confer additional risk in the presence of *BRCA2* but not *BRCA1* mutations (12). This interesting disparity may in part reflect the association of the SNPs with ER-positive tumors because *BRCA1* is typically associated with ER-negative tumors. In a recent candidate-based experiment, it was reported that homozygosity for a variant in the 5′ untranslated region of *RAD51* confers an increased risk of breast cancer to *BRCA2* mutation carriers [hazard ratio = 3.18 (95% CI = 1.39–7.27)]. The modifying effect was not significant in *BRCA1* mutation carriers or when only a single copy of the risk allele was present (11). By contrast, *CHEK2*_1100delC has not been shown to confer an elevated risk on the background of mutations in either *BRCA1* or *BRCA2* (85). It has been proposed that this reflects the common pathway of the encoded proteins; abrogation of CHK2 function might have little additional impact on a pathway already radically subverted by a mutation in *BRCA1* or *BRCA2*. However, there is currently no biological proof of this hypothesis. Further analyses are required to establish whether similar interactions are observed for *ATM*, *BRIP1*, and *PALB2*, which also interact with BRCA1 and/or BRCA2 in DNA repair pathways.

Clarity regarding gene-environment interactions is even more limited. Recognized environmental risk factors for breast cancer in the general population are well established and predominantly relate to estrogen exposure. Endogenous estrogen-related risk factors include timing of menarche and menopause,

parity, age of first live birth, and breast-feeding, whereas exogenous factors include administration of contraceptives and hormone replacement therapy (30, 31, 33). Studies of hormonal/reproductive factors in *BRCA1* and *BRCA2* mutation carriers are challenging, not just in the assembly of sufficient families, but because of the biases inherent in observation of the behaviors of a group of individuals who know themselves to be at elevated risk. However, well-powered studies of gene-environment interactions are becoming possible through collaboration and are allowing the effects of *BRCA1* and *BRCA2* to be studied independently. Significant alteration of breast cancer risk in *BRCA1* and *BRCA2* mutation carriers has been demonstrated for oral contraceptive usage, age at first pregnancy, and degree of parity, whereas significant effects were not demonstrated for age of menarche, age of natural menopause, nulliparity, and breast feeding (3, 23, 25). These observations warrant further investigation and confirmation in prospective studies. Other large studies will be required to explore gene-environment interactions for genetic predisposition factors of lower penetrance.

## IDENTIFICATION OF FURTHER GENETIC BREAST CANCER PREDISPOSITION FACTORS

There is a clear discontinuity in the risks associated with the three categories of breast cancer predisposition factors identified to date: The high-penetrance genes confer a risk that is elevated >10-fold, the known intermediate-penetrance genes 2–4-fold, and the low-penetrance alleles <1.5-fold. There may be some biological significance to these strata or they may be an artifact of the limited methods of ascertainment. More than 70% of genetic predisposition to breast cancer remains unaccounted for and as further genetic predisposition factors are identified, new categories may emerge and/or the apparent distinctions between extant classes may blur or disappear. Nevertheless, risk and prevalence provide a useful framework for considering how technological and intellectual advances may allow extension of the repertoire of breast cancer predisposition factors.

### High-Penetrance Genes

Any further high-penetrance dominant predisposition genes are likely to be very rare causes of familial breast cancer. Segregation analyses generated no evidence for further dominant genes of a risk-penetrance profile comparable to *BRCA1* or *BRCA2* and this has been corroborated by linkage studies. In a recently published example, genome-wide linkage analysis was undertaken in 149 nonsyndromic breast cancer families negative for mutations in *BRCA1* and *BRCA2* and the number of linkage peaks detected under parametric (dominant and recessive) and nonparametric (allele-sharing) models did not differ significantly from that expected by chance (109). However, linkage analyses such as this have been undertaken in families largely ascertained on the basis of multiple-generation breast cancer pedigrees. A highly penetrant recessive predisposition gene would not produce this pattern of disease in families. Two independent segregation analyses found evidence of a recessive pattern of inheritance (8, 35). Thus, there may be utility in genome-wide recessive linkage analyses in more suitable series, in particular families from understudied population isolates, and/or in families with higher levels of consanguinity. Further, unidentified breast cancer-associated syndromes may also exist. However, if these syndromes have not yet come to medical/scientific attention, they are likely to be very rare; linkage is likely to represent the optimal method for mapping any such underlying predisposition genes.

### Intermediate-Penetrance Genes

Further intermediate-penetrance breast cancer predisposition genes likely exist, although it is difficult to predict how many there may be and what proportion of the total excess familial risk is attributable to them. Resequencing

is currently the optimal approach by which to identify genes of this nature. The known intermediate-penetrance genes are all involved in DNA repair and function in pathways with BRCA1 and/or BRCA2; this commonality may reflect the underlying biology of this class or may just be an artifact of the groups of genes that have been most intensely investigated. Further genes involved in DNA repair and other relevant pathways represent plausible candidates and are being investigated by us and other groups. However, technological advances are facilitating increasingly high-throughput mutational screening such that genome-wide resequencing is becoming viable. This will allow interrogation of regions of the genome not previously investigated and may reveal further intermediate-penetrance genes that function in pathways not predictable from current paradigms.

Intermediate-penetrance predisposition to breast cancer may also be an unrecognized component of known pleomorphic cancer predisposition syndromes. Large, collaborative epidemiological studies of rare syndromes can optimize power and minimize bias and may represent the optimal strategy by which to establish accurate estimates of these breast cancer risks. Epidemiological studies of recessive syndromes, particularly those associated with childhood cancer, may reveal elevated frequency of breast cancers in relatives of affected individuals and may lead to the identification of further intermediate-penetrance breast cancer genes, as in the case of *ATM*.

## Low-Penetrance Variants

**Common low-penetrance variants.** Extant genome-wide association data offer clear evidence that many further common low-penetrance breast cancer predisposition variants exist. Further risk alleles may be identified from these data by mining the variants of borderline significance for further signals and through the use of imputation techniques. However, comprehensive study of this (likely) extensive repertoire of low-penetrance alleles will require further, larger, genome-wide association experiments. The effect sizes of subsequent rounds of risk alleles are likely to be of progressively diminishing magnitude and will require commensurate increases in power for detection. It is currently unclear what proportion of common low-penetrance alleles will be detectable by the feasible studies in the immediate future.

Strongly associated tag SNPs from genome-wide scans may be utilized for genetic epidemiologic analyses and clinical risk estimation in their own right. However, the way in which these common variants contribute to cancer is largely unknown and exploration of the biological mechanisms that underlie these signals offers exciting new avenues of study. Resequencing and fine-association mapping of the region tagged by a reporter SNP can refine the association to a minimum set of SNPs in a fixed block of linkage disequilibrium. The cancer-causing components within these blocks are cryptic. Attempts to identify them may challenge current paradigms of the relationship between genes and disease and require innovative methods.

To date, genome-wide studies have been undertaken in simple series of breast cancer cases of European descent. Novel risk alleles may be identifiable through genome-wide studies of different ethnic groups in which the minor allele frequencies and phenotypic spectrum of disease differ from Europeans. For example, a risk locus at 6q22 was identified in a recent genome-wide study performed exclusively in Ashkenazi Jewish breast cancer cases and controls; verification of this association in other populations is awaited (54). Broader insight into genetic etiology may be gained through the expansion of genome-wide studies into subgroups of breast cancer cases. For example, analysis of *BRCA*-positive individuals may allow identification of modifier alleles. Studies focusing on well-characterized histopathological subgroups of breast cancers may advance our understanding of the genetic basis of disease heterogeneity. The use of quantitative intermediate phenotypes, such as mammographic density, as the

outcome for genome-wide studies may also represent an alternative approach.

**Rare low-penetrance variants.** The risk-prevalence profiles already identified suggest that rare variants of low penetrance represent another plausible category of breast cancer predisposition factor. However, there are currently no reliable methods by which to quantify the risk of breast cancer associated with individual, rare variants. Therefore, although it is possible that some of the rare variants detected during mutational screening experiments are low-penetrance predisposition factors, our ability to demonstrate this is limited. Genome-wide resequencing will inevitably detect many further rare variants, some of which will likely be associated with small increases in risk of breast cancer. Developing robust methods for the identification of the cancer-associated rare variants among the numerous innocuous variants will be a major challenge for the future.

## Optimizing the Power of Gene Identification Studies

A priority for all future studies will likely be the harnessing of sufficient statistical power at acceptable cost. An important tool for gaining power without increasing the experiment size/cost is to assemble a case series enriched for genetic predisposition. The best-recognized of these enrichment parameters, successfully used in many experiments to date, include family history, bilaterality, and early age of onset of disease. Their relative efficacies have been compared through modeling the sample size required to demonstrate association with disease of a variant of specified frequency and effect. The use of cases with a single affected first-degree relative affords a greater than twofold reduction in the required sample size; for cases with two affected first-degree relatives the reduction is more than fourfold. The sample size reduction when using bilateral cases is also fourfold (and thus equivalent to using cases with two first-degree relatives). However, the sample size required for an association study that uses cases diagnosed at age 35 is only 40% less than one that uses cases diagnosed at 65 years. Relative efficiencies are all magnified for experiments in which rarer alleles are studied (6).

An alternative strategy is geographical enrichment, namely to first screen genes in population isolates in which a higher prevalence of founder mutations might be anticipated, such as the Ashkenazim, the Finnish, or the Icelanders. This has proved very successful for the identification and characterization of known genes. However, as is currently the case for *RAD50*, the generalizability of findings from specific populations may prove challenging.

There has been longstanding expectation that evolution in molecular profiling may result in pathological classifications that more directly reflect the genetic etiology of tumors. Although the basal phenotype clearly enriches for cases that arise because of *BRCA1* mutations, it remains unclear whether novel immunohistopathological profiles or other phenotypic surrogates may emerge that can distinguish further subsets of breast cancers that occur because of genetic predisposition.

# CLINICAL TRANSLATION

## Risk Estimation and Management

Risk estimation is currently the primary clinical application for genetic factors that predispose to breast cancer. Ongoing improvement of clinical risk assessment tools is necessary to ensure that individuals at truly elevated risk are identified so that they might benefit from advances in surveillance techniques and prophylactic interventions. Advances in the understanding of the polygenic basis of breast cancer may have different effects on three groups: families positive for mutations in *BRCA1* or *BRCA2*, other breast cancer families, and the general population.

**Risk estimation in BRCA-positive breast cancer families.** Detection in a family of a mutation in *BRCA1* or *BRCA2* has afforded relative clarity in risk estimation. Unaffected females,

at high prior risk of cancer on account of their family history, can undertake predictive testing to determine whether they carry the high-risk mutation. Risk-reducing prophylactic surgery and/or intensive surveillance may be offered to mutation carriers whereas noncarriers can be reassured that their risk is not as high. However, studies have found wide variation in the penetrance of mutations in *BRCA1* and *BRCA2* (see above) and this has raised questions regarding the appropriate estimates of risk for clinical usage. Some of this apparent variation in penetrance between families may be the result of differing doses of additional lower-risk modifying variants; recently discovered examples include SNPs in *RAD51*, *FGFR2*, and at 5q (11, 12). Thus, one of the early clinical applications of low-penetrance alleles may be to offer individualized refinement of risk to *BRCA1* and *BRCA2* mutation carriers.

**Risk estimation in BRCA-negative breast cancer families.** The majority of breast cancer families do not harbor mutations in *BRCA1* or *BRCA2*. Clinical breast cancer risk estimation in these families is currently based empirically upon family history of cancer. For unaffected individuals in breast cancer families negative for *BRCA1* and *BRCA2* mutations, it would be clinically useful to have better discrimination of risk to distinguish high-risk individuals in whom radical surgical prophylaxis may be justified and individuals at lower/population risk who could be spared some anxiety and intervention. It remains unclear whether the underlying genetic architecture is such that this discrimination could be provided by genotyping multiple genetic predisposition factors (presuming sufficient numbers were identified). If the clustering of breast cancer in a family has occurred because of multiple low-penetrance genetic factors, typing of these factors would result in many possible risk categories. Genotyping may place the majority of unaffected members of the family into intermediate risk categories that differ little from their risk based on family history. If family clustering is the result of fewer factors of greater penetrance, genotyping these may adjust risk estimation sufficiently to alter clinical management for unaffected individuals.

**Risk estimation in the general population.** By contrast, in the absence of a family history of breast cancer, genotyping is the only means of identifying individuals at increased genetic risk. The discrimination of risk that could be afforded by population-level genotyping of common low-penetrance variants is influenced by two factors: firstly, the true extent to which breast cancer risk varies across the population (which is currently unclear) and secondly, the proportion of the risk variants that is available for genotyping (8, 9, 94). Although the logistical, ethical, social, and economic considerations are numerous, in principle it seems plausible that population-based genotyping could eventually be used to aid stratification and resource allocation. Surveillance, primary antiestrogen chemoprophylaxis, prophylactic surgery, and other emerging interventions may be appropriate for the high-risk upper tail of the population. The low-risk tail may require little or no additional intervention.

## Targeted Therapies

Understanding of the biological mechanisms underlying breast cancer predisposition genes is beginning to offer exciting opportunities for new therapies. The roles of BRCA1 and BRCA2 in DNA repair and the resultant sensitivity of BRCA1- and BRCA2-deficient cells to agents that cross-link DNA are being exploited by studies in which mutation carriers are treated with platinum-based drugs. The inherent DNA repair defect in BRCA-deficient cells has also provided a rationale for a further therapeutic approach. Poly (ADP-ribose) polymerase (PARP) is an enzyme involved in base excision repair. Inhibition of PARP results in an increase in DNA lesions that are normally repaired through homologous recombination, which requires BRCA1 and BRCA2. In a background deficient for either BRCA1 or BRCA2 protein, cells are profoundly sensitive to

inhibition by PARP, which results in cell cycle arrest, chromosome instability, and cell death. Thus in *BRCA* mutation carriers, PARP inhibitors are synthetically lethal to tumor cells but confer no demonstrable toxicity to normal heterozygous cells. Studies are also underway to investigate whether PARP inhibitors have similar effects in individuals with basal tumors similar to those occurring in *BRCA1* mutation carriers and/or cancers associated with defects in other DNA repair proteins (50, 84). Understanding the molecular basis of the increased cancer risk associated with low-penetrance alleles is in its infancy. However, this understanding may offer opportunities for novel therapies. For example, if some risk alleles drive tumorigenesis through the upregulation of oncogenes, such as *FGFR2* and/or *MYC*, these genes may represent potential new targets for therapeutic interventions.

## CONCLUSION

Although recent breakthroughs have resulted in the identification of distinct new groups of genetic breast cancer predisposition factors, more than 70% of the genetic predisposition to breast cancer remains unexplained. Technologies are emerging rapidly that may allow us to identify many further predisposition factors within these classes and of other risk-prevalence profiles. The identification of novel predisposition factors offers exciting challenges, but ongoing clarification and characterization of known genetic risk factors is also important. A formidable and ongoing challenge is to marshal disparate experimental results to capture, validate, qualify, and organize into an integrated schema all the emerging components of risk. This is particularly important if we are to advance clinical risk estimation to optimally apportion surveillance and prophylactic interventions.

### SUMMARY POINTS

1. Multiple strategies have been used to identify breast cancer–predisposing genetic factors.
2. Linkage analysis has been successful in mapping high-penetrance genes such as *BRCA1* and *BRCA2*.
3. Mutational screening of candidate genes has been used to identify genes such as *CHEK2*, *ATM*, *BRIP1*, and *PALB2*, mutations in which are rare and confer intermediate penetrance of breast cancer.
4. Genome-wide association studies have revealed several low-penetrance alleles.
5. More than 70% of breast cancer predisposition remains unexplained.
6. The outstanding predisposition is likely to be polygenic and may include multiple further common low-penetrance alleles, rare intermediate-penetrance genes, and rare low-penetrance alleles.
7. *BRCA1* and *BRCA2* mutation testing allows identification of individuals at elevated risk of breast cancer who can be offered risk-reducing interventions.
8. Targeted therapies are being developed that exploit the biological functions of BRCA1 and BRCA2.

### FUTURE ISSUES

1. Further genome-wide association studies are required to identify further common variants.

2. Genome-wide resequencing is likely to detect numerous novel rare variants, some of which may predispose to breast cancer.
3. Understanding how novel variants result in breast cancer predisposition may require innovative strategies.
4. Quantification of gene-gene interactions and gene-environment interactions, which may be heterogeneous in nature, may be used to improve risk estimation.
5. Judicious clinical translation of genetic factors in addition to *BRCA1* and *BRCA2* may assist in risk estimation, optimization of management, and development of therapies.

## DISCLOSURE STATEMENT

The authors are not aware of any biases that might be perceived as affecting the objectivity of this review.

## ACKNOWLEDGMENTS

We are grateful to Mike Stratton, Anthony Renwick, and Richard Scott for their critical reading of the manuscript and to Peter Donnelly, Julian Maller, and Paul Pharoah for their assistance in analyses of the low penetrance alleles.

## LITERATURE CITED

1. Alessi DR, Sakamoto K, Bayascas JR. 2006. LKB1-dependent signaling pathways. *Annu. Rev. Biochem.* 75:137–63
2. Allinen M, Huusko P, Mantyniemi S, Launonen V, Winqvist R. 2001. Mutation analysis of the *CHK2* gene in families with hereditary breast cancer. *Br. J. Cancer* 85:209–12
3. Andrieu N, Goldgar DE, Easton DF, Rookus M, Brohet R, et al. 2006. Pregnancies, breast-feeding, and breast cancer risk in the International BRCA1/2 Carrier Cohort Study (IBCCS). *J. Natl. Cancer Inst.* 98:535–44
4. Anglian Breast Cancer Study Group. 2000. Prevalence and penetrance of *BRCA1* and *BRCA2* mutations in a population-based series of breast cancer cases. *Br. J. Cancer* 83:1301–8
5. Antoniou A, Pharoah PD, Narod S, Risch HA, Eyfjord JE, et al. 2003. Average risks of breast and ovarian cancer associated with *BRCA1* or *BRCA2* mutations detected in case series unselected for family history: a combined analysis of 22 studies. *Am. J. Hum. Genet.* 72:1117–30
6. Antoniou AC, Easton DF. 2003. Polygenic inheritance of breast cancer: Implications for design of association studies. *Genet. Epidemiol.* 25:190–202
7. Antoniou AC, Gayther SA, Stratton JF, Ponder BA, Easton DF. 2000. Risk models for familial ovarian and breast cancer. *Genet. Epidemiol.* 18:173–90
8. Antoniou AC, Pharoah PD, McMullan G, Day NE, Ponder BA, Easton D. 2001. Evidence for further breast cancer susceptibility genes in addition to *BRCA1* and *BRCA2* in a population-based study. *Genet. Epidemiol.* 21:1–18
9. Antoniou AC, Pharoah PD, McMullan G, Day NE, Stratton MR, et al. 2002. A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. *Br. J. Cancer* 86:76–83
10. Antoniou AC, Pharoah PP, Smith P, Easton DF. 2004. The BOADICEA model of genetic susceptibility to breast and ovarian cancer. *Br. J. Cancer* 91:1580–90

11. Antoniou AC, Sinilnikova OM, Simard J, Leone M, Dumont M, et al. 2007. *RAD51* 135G–>C modifies breast cancer risk among *BRCA2* mutation carriers: results from a combined analysis of 19 studies. *Am. J. Hum. Genet.* 81:1186–200
12. Antoniou AC, Spurdle AB, Sinilnikova OM, Healey S, Pooley KA, et al. 2008. Common breast cancer-predisposition alleles are associated with breast cancer risk in BRCA1 and BRCA2 mutation carriers. *Am. J. Hum. Genet.* 82:937–48
13. Bell DW, Varley JM, Szydlo TE, Kang DH, Wahrer DC, et al. 1999. Heterozygous germ line *hCHK2* mutations in Li-Fraumeni syndrome. *Science* 286:2528–31
14. Bignell GR, Barfoot R, Seal S, Collins N, Warren W, Stratton MR. 1998. Low frequency of somatic mutations in the *LKB1*/Peutz-Jeghers syndrome gene in sporadic breast cancer. *Cancer Res.* 58:1384–86
15. Birch JM, Alston RD, McNally RJ, Evans DG, Kelsey AM, et al. 2001. Relative frequency and morphology of cancers in carriers of germline TP53 mutations. *Oncogene* 20:4621–28
16. Bishop DT, Cannon-Albright L, McLellan T, Gardner EJ, Skolnick MH. 1988. Segregation and linkage analysis of nine Utah breast cancer pedigrees. *Genet. Epidemiol.* 5:151–69
17. Bogdanova N, Enssen-Dubrowinskaja N, Feshchenko S, Lazjuk GI, Rogov YI, et al. 2005. Association of two mutations in the *CHEK2* gene with breast cancer. *Int. J. Cancer* 116:263–66
18. Borresen AL, Andersen TI, Garber J, Barbier-Piraux N, Thorlacius S, et al. 1992. Screening for germ line *TP53* mutations in breast cancer patients. *Cancer Res.* 52:3234–36
19. Breast Cancer Assoc. Consort. 2006. Commonly studied single-nucleotide polymorphisms and breast cancer: results from the Breast Cancer Association Consortium. *J. Natl. Cancer Inst.* 98:1382–96
20. Breast Cancer Link. Consort. 1997. Pathology of familial breast cancer: differences between breast cancers in carriers of BRCA1 or BRCA2 mutations and sporadic cases. *Lancet* 349:1505–10
21. Breast Cancer Link. Consort. 1999. Cancer risks in BRCA2 mutation carriers. *J. Natl. Cancer Inst.* 91:1310–16
22. Broca P. 1866. *Traité des Tumeurs*. Paris: Asselin
23. Brohet RM, Goldgar DE, Easton DF, Antoniou AC, Andrieu N, et al. 2007. Oral contraceptives and breast cancer risk in the international BRCA1/2 carrier cohort study: a report from EMBRACE, GENEPSO, GEO-HEBON, and the IBCCS Collaborating Group. *J. Clin. Oncol.* 25:3831–36
24. Cantor SB, Bell DW, Ganesan S, Kass EM, Drapkin R, et al. 2001. BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. *Cell* 105:149–60
25. Chang-Claude J, Andrieu N, Rookus M, Brohet R, Antoniou AC, et al. 2007. Age at menarche and menopause and breast cancer risk in the International *BRCA1/2* Carrier Cohort Study. *Cancer Epidemiol. Biomark. Prev.* 16:740–46
26. CHEK2 Breast Cancer Case-Control Consort. 2004. CHEK2*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from 10 studies. 2004. *Am. J. Hum. Genet.* 74:1175–82 (WAS 8)
27. Chen J, Lindblom P, Lindblom A. 1998. A study of the PTEN/MMAC1 gene in 136 breast cancer families. *Hum. Genet.* 102:124–25
28. Chompret A, Brugieres L, Ronsin M, Gardes M, Dessarps-Freichey F, et al. 2000. p53 germline mutations in childhood cancers and cancer risk for carrier individuals. *Br. J. Cancer* 82:1932–37
29. Claus EB, Risch N, Thompson WD. 1991. Genetic analysis of breast cancer in the cancer and steroid hormone study. *Am. J. Hum. Genet.* 48:232–42
30. Collab. Group Horm. Factors Breast Cancer. 1996. Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53,297 women with breast cancer and 100,239 women without breast cancer from 54 epidemiological studies. *Lancet* 347:1713–27
31. Collab. Group Horm. Factors Breast Cancer. 1997. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. *Lancet* 350:1047–59
32. Collab. Group Horm. Factors Breast Cancer. 2001. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet* 358:1389–99

33. Collab. Group Horm. Factors Breast Cancer. 2002. Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50,302 women with breast cancer and 96,973 women without the disease. *Lancet* 360:187–95
34. Cox A, Dunning AM, Garcia-Closas M, Balasubramanian S, Reed MW, et al. 2007. A common coding variant in *CASP8* is associated with breast cancer risk. *Nat. Genet.* 39:352–58
35. Cui J, Antoniou AC, Dite GS, Southey MC, Venter DJ, et al. 2001. After BRCA1 and BRCA2-what next? Multifactorial segregation analyses of three-generation, population-based Australian families affected by female breast cancer. *Am. J. Hum. Genet.* 68:420–31
36. Cully M, You H, Levine AJ, Mak TW. 2006. Beyond PTEN mutations: the PI3K pathway as an integrator of multiple inputs during tumorigenesis. *Nat. Rev. Cancer* 6:184–92
37. Cybulski C, Górski B, Huzarski T, Masojć B, Mierzejewski M, et al. 2004. *CHEK2* is a multiorgan cancer susceptibility gene. *Am. J. Hum. Genet.* 75:1131–35
38. Dong X, Wang L, Taniguchi K, Wang X, Cunningham JM, et al. 2003. Mutations in *CHEK2* associated with prostate cancer risk. *Am. J. Hum. Genet.* 72:270–80
39. Dufault MR, Betz B, Wappenschmidt B, Hofmann W, Bandick K, et al. 2004. Limited relevance of the *CHEK2* gene in hereditary breast cancer. *Int. J. Cancer* 110:320–25
40. Dunning AM, Healey CS, Pharoah PD, Teare MD, Ponder BA, Easton DF. 1999. A systematic review of genetic polymorphisms and breast cancer risk. *Cancer Epidemiol. Biomark. Prev.* 8:843–54
41. Easton DF. 1994. Cancer risks in A-T heterozygotes. *Int. J. Radiat. Biol.* 66:S177–82
42. Easton DF. 1999. How many more breast cancer predisposition genes are there? *Breast Cancer Res.* 1:14–47
43. Easton DF, Deffenbaugh AM, Pruss D, Frye C, Wenstrup RJ, et al. 2007. A systematic genetic assessment of 1433 sequence variants of unknown clinical significance in the BRCA1 and BRCA2 breast cancer-predisposition genes. *Am. J. Hum. Genet.* 81:873–83
44. Easton DF, Ford D, Bishop DT. 1995. Breast and ovarian cancer incidence in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. *Am. J. Hum. Genet.* 56:265–71
45. Easton DF, Matthews FE, Ford D, Swerdlow AJ, Peto J. 1996. Cancer mortality in relatives of women with ovarian cancer: the OPCS Study. Office of Population Censuses and Surveys. *Int. J. Cancer* 65:284–94
46. Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, et al. 2007. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 447:1087–93
47. Erkko H, Xia B, Nikkila J, Schleutker J, Syrjakoski K, et al. 2007. A recurrent mutation in *PALB2* in Finnish cancer families. *Nature* 446:316–19
48. Evans DG, Birch JM, Thorneycroft M, McGown G, Lalloo F, Varley JM. 2002. Low rate of *TP53* germline mutations in breast cancer/sarcoma families not fulfilling classical criteria for Li-Fraumeni syndrome. *J. Med. Genet.* 39:941–44
49. Evers B, Jonkers J. 2006. Mouse models of BRCA1 and BRCA2 deficiency: past lessons, current understanding and future prospects. *Oncogene* 25:5885–97
50. Farmer H, McCabe N, Lord CJ, Tutt AN, Johnson DA, et al. 2005. Targeting the DNA repair defect in *BRCA* mutant cells as a therapeutic strategy. *Nature* 434:917–21
51. Ford D, Easton DF, Stratton M, Narod S, Goldgar D, et al. 1998. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am. J. Hum. Genet.* 62:676–89
52. Foulkes WD, Ghadirian P, Akbari MR, Hamel N, Giroux S, et al. 2007. Identification of a novel truncating *PALB2* mutation and analysis of its contribution to early-onset breast cancer in French Canadian women. *Breast Cancer Res.* 9:R83
53. Garcia-Closas M, Hall P, Nevanlinna H, Pooley K, Morrison J, et al. 2008. Heterogeneity of breast cancer associations with five susceptibility loci by clinical and pathological characteristics. *PLoS Genet.* 4:e1000054
54. Gold B, Kirchhoff T, Stefanov S, Lautenberger J, Viale A, et al. 2008. Genome-wide association study provides evidence for a breast cancer risk locus at 6q22.33. *Proc. Natl. Acad Sci. USA* 105:4340–5
55. Górski B, Cybulski C, Huzarski T, Byrski T, Gronwald J, et al. 2005. Breast cancer predisposing alleles in Poland. *Breast Cancer Res. Treat.* 92:19–24

56. Gudmundsdottir K, Ashworth A. 2006. The roles of BRCA1 and BRCA2 and associated proteins in the maintenance of genomic stability. *Oncogene* 25:5864–74
57. Guénard F, Labrie Y, Ouellette G, Beauparlant CJ, Bessette P, et al. 2007. Germline mutations in the breast cancer susceptibility gene *PTEN* are rare in high-risk non-*BRCA1/2* French Canadian breast cancer families. *Fam. Cancer* 6:483–90
58. Guilford P, Hopkins J, Harraway J, McLeod M, McLeod N, et al. 1998. E-cadherin germline mutations in familial gastric cancer. *Nature* 392:402–5
59. Haiman CA, Le ML, Yamamato J, Stram DO, Sheng X, et al. 2007. A common genetic risk factor for colorectal and prostate cancer. *Nat. Genet.* 39:954–56
60. Haiman CA, Patterson N, Freedman ML, Myers SR, Pike MC, et al. 2007. Multiple regions within 8q24 independently affect risk for prostate cancer. *Nat. Genet.* 39:638–44
61. Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, et al. 1990. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* 250:1684–89
62. Heikkinen K, Rapakko K, Karppinen SM, Erkko H, Knuutila S, et al. 2006. *RAD50* and *NBS1* are breast cancer susceptibility genes associated with genomic instability. *Carcinogenesis* 27:1593–99
63. Hemminki A, Markie D, Tomlinson I, Avizienyte E, Roth S, et al. 1998. A serine/threonine kinase gene defective in Peutz-Jeghers syndrome. *Nature* 391:184–87
64. Hemminki A, Tomlinson I, Markie D, Jarvinen H, Sistonen P, et al. 1997. Localization of a susceptibility locus for Peutz-Jeghers syndrome to 19p using comparative genomic hybridization and targeted linkage analysis. *Nat. Genet.* 15:87–90
65. Hopper JL, Carlin JB. 1992. Familial aggregation of a disease consequent upon correlation between relatives in a risk factor measured on a continuous scale. *Am. J. Epidemiol.* 136:1138–47
66. Howlett NG, Taniguchi T, Olson S, Cox B, Waisfisz Q, et al. 2002. Biallelic inactivation of *BRCA2* in Fanconi anemia. *Science* 297:606–9
67. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, et al. 2007. A genome-wide association study identifies alleles in *FGFR2* associated with risk of sporadic postmenopausal breast cancer. *Nat. Genet.* 39:870–74
68. Iselius L, Slack J, Littler M, Morton NE. 1991. Genetic epidemiology of breast cancer in Britain. *Ann. Hum. Genet.* 55:151–59
69. Jacquemier J, Padovani L, Rabayrol L, Lakhani SR, Penault-Llorca F, et al. 2005. Typical medullary breast carcinomas have a basal/myoepithelial phenotype. *J. Pathol.* 207:260–68
70. Jenne DE, Reimann H, Nezu J, Friedel W, Loff S, et al. 1998. Peutz-Jeghers syndrome is caused by mutations in a novel serine threonine kinase. *Nat. Genet.* 18:38–43
71. Keller G, Vogelsang H, Becker I, Hutter J, Ott K, et al. 1999. Diffuse type gastric and lobular breast carcinoma in a familial gastric cancer patient with an E-cadherin germline mutation. *Am. J. Pathol.* 155:337–42
72. Kilpivaara O, Vahteristo P, Falck J, Syrjakoski K, Eerola H, et al. 2004. CHEK2 variant I157T may be associated with increased breast cancer risk. *Int. J. Cancer* 111:543–47
73. Lakhani SR. 1999. The pathology of familial breast cancer: Morphological aspects. *Breast Cancer Res.* 1:31–35
74. Lakhani SR, Reis-Filho JS, Fulford L, Penault-Llorca F, van der Vijver M, et al. 2005. Prediction of *BRCA1* status in patients with breast cancer using estrogen receptor and basal phenotype. *Clin. Cancer Res.* 11:5175–80
75. Lakhani SR, van der Vijver M, Jacquemier J, Anderson TJ, Osin PP, et al. 2002. The pathology of familial breast cancer: predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in *BRCA1* and *BRCA2*. *J. Clin. Oncol.* 20:2310–18
76. Lalloo F, Varley J, Moran A, Ellis D, O'dair L, et al. 2006. *BRCA1*, *BRCA2* and *TP53* mutations in very early-onset breast cancer with associated risks to relatives. *Eur. J. Cancer* 42:1143–50
77. Le Dran H. 1757. Memoire avec un precis de plusieurs observations sur le cancer. *Mem. Acad. R. Chir.* 3:1–54
78. Levitus M, Waisfisz Q, Godthelp BC, de Vries Y, Hussain S, et al. 2005. The DNA helicase BRIP1 is defective in Fanconi anemia complementation group. *J. Nat. Genet.* 37:934–35

79. Levran O, Attwooll C, Henry RT, Milton KL, Neveling K, et al. 2005. The BRCA1-interacting helicase BRIP1 is deficient in Fanconi anemia. *Nat. Genet.* 37:931–33
80. Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, et al. 2000. Environmental and heritable factors in the causation of cancer–analyses of cohorts of twins from Sweden, Denmark, and Finland. *N. Engl. J. Med.* 343:78–85
81. Litman R, Peng M, Jin Z, Zhang F, Zhang J, et al. 2005. BACH1 is critical for homologous recombination and appears to be the Fanconi anemia gene product FANCJ. *Cancer Cell* 8:255–65
82. Malkin D, Li FP, Strong LC, Fraumeni JF Jr, Nelson CE, et al. 1990. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science* 250:1233–38
83. Masciari S, Larsson N, Senz J, Boyd N, Kaurah P, et al. 2007. Germline E-cadherin mutations in familial lobular breast cancer. *J. Med. Genet.* 44:726–31
84. McCabe N, Turner NC, Lord CJ, Kluzek K, Bialkowska A, et al. 2006. Deficiency in the repair of DNA damage by homologous recombination and sensitivity to poly(ADP-ribose) polymerase inhibition. *Cancer Res.* 66:8109–15
85. Meijers-Heijboer H, van den Ouweland A, Klijn J, Wasielewski M, de Snoo A, et al. 2002. Low-penetrance susceptibility to breast cancer due to *CHEK2**1100delC in noncarriers of *BRCA1* or *BRCA2* mutations. *Nat. Genet.* 31:55–59
86. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, et al. 1994. A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. *Science* 266:66–71
87. Nelen MR, Padberg GW, Peeters EA, Lin AY, van den Helm B, et al. 1996. Localization of the gene for Cowden disease to chromosome 10q22–23. *Nat. Genet.* 13:114–16
88. Nelen MR, van Staveren WC, Peeters EA, Hassel MB, Gorlin RJ, et al. 1997. Germline mutations in the *PTEN/MMAC1* gene in patients with Cowden disease. *Hum. Mol. Genet.* 6:1383–87
89. Parmigiani G, Berry D, Aguilar O. 1998. Determining carrier probabilities for breast cancer-susceptibility genes BRCA1 and BRCA2. *Am. J. Hum. Genet.* 62:145–58
90. Peng M, Litman R, Jin Z, Fong G, Cantor SB. 2006. BACH1 is a DNA repair protein supporting BRCA1 damage response. *Oncogene* 25:2245–53
91. Peto J, Collins N, Barfoot R, Seal S, Warren W, et al. 1999. Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer. *J. Natl. Cancer Inst.* 91:943–49
92. Peto J, Easton DF, Matthews FE, Ford D, Swerdlow AJ. 1996. Cancer mortality in relatives of women with breast cancer: the OPCS Study. Office of Population Censuses and Surveys. *Int. J. Cancer* 65:275–83
93. Peto J, Mack TM. 2000. High constant incidence in twins and other relatives of women with breast cancer. *Nat. Genet.* 26:411–14
94. Pharoah PD, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BA. 2002. Polygenic susceptibility to breast cancer and implications for prevention. *Nat. Genet.* 31:33–36
95. Pharoah PD, Dunning AM, Ponder BA, Easton DF. 2004. Association studies for finding cancer-susceptibility genetic variants. *Nat. Rev. Cancer* 4:850–60
96. Pharoah PD, Guilford P, Caldas C. 2001. Incidence of gastric cancer and breast cancer in CDH1 (E-cadherin) mutation carriers from hereditary diffuse gastric cancer families. *Gastroenterology* 121:1348–53
97. Rahman N, Scott RH. 2007. Cancer genes associated with phenotypes in monoallelic and biallelic mutation carriers: new lessons from old players. *Hum. Mol. Genet.* 16(Spec. No. 1):R60–66
98. Rahman N, Seal S, Thompson D, Kelly P, Renwick A, et al. 2007. *PALB2*, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat. Genet.* 39:165–67
99. Ramus SJ, Harrington PA, Pye C, Dicioccio RA, Cox MJ, et al. 2007. Contribution of *BRCA1* and *BRCA2* mutations to inherited ovarian cancer. *Hum. Mutat.* 28:1207–15
100. Reid S, Schindler D, Hanenberg H, Barker K, Hanks S, et al. 2007. Biallelic mutations in *PALB2* cause Fanconi anemia subtype FA-N and predispose to childhood cancer. *Nat. Genet.* 39:162–64
101. Renwick A, Thompson D, Seal S, Kelly P, Chagtai T, et al. 2006. *ATM* mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. *Nat. Genet.* 38:873–75
102. Savitsky K, Bar-Shira A, Gilad S, Rotman G, Ziv Y, et al. 1995. A single ataxia telangiectasia gene with a product similar to PI-3 kinase. *Science* 268:1749–53

103. Schutte M, Seal S, Barfoot R, Meijers-Heijboer H, Wasielewski M, et al. 2003. Variants in *CHEK2* other than 1100delC do not make a major contribution to breast cancer susceptibility. *Am. J. Hum. Genet.* 72:1023–28
104. Seal S, Thompson D, Renwick A, Elliott A, Kelly P, et al. 2006. Truncating mutations in the Fanconi anemia J gene *BRIP1* are low-penetrance breast cancer susceptibility alleles. *Nat. Genet.* 38:1239–41
105. Seppala EH, Ikonen T, Mononen N, Autio V, Rokman A, et al. 2003. *CHEK2* variants associate with hereditary prostate cancer. *Br. J. Cancer* 89:1966–70
106. Sharan SK, Wims M, Bradley A. 1995. Murine *Brca1*: sequence and significance for human missense mutations. *Hum. Mol. Genet.* 4:2275–78
107. Shugart YY, Cour C, Renard H, Lenoir G, Goldgar D, et al. 1999. Linkage analysis of 56 multiplex families excludes the Cowden disease gene *PTEN* as a major contributor to familial breast cancer. *J. Med. Genet.* 36:720–21
108. Smid M, Wang Y, Klijn JG, Sieuwerts AM, Zhang Y, et al. 2006. Genes associated with breast cancer metastatic to bone. *J. Clin. Oncol.* 24:2261–67
109. Smith P, McGuffog L, Easton DF, Mann GJ, Pupo GM, et al. 2006. A genome wide linkage search for breast cancer susceptibility genes. *Genes Chromosomes Cancer* 45:646–55
110. Stacey SN, Manolescu A, Sulem P, Rafnar T, Gudmundsson J, et al. 2007. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat. Genet.* 39:865–69
111. Stacey SN, Manolescu A, Sulem P, Thorlacius S, Gudjonsson SA, et al. 2008. Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat. Genet.* doi:10.1038/ng.131
112. Swift M, Sholman L, Perry M, Chase C. 1976. Malignant neoplasms in the families of patients with ataxia-telangiectasia. *Cancer Res.* 36:209–15
113. Thompson D, Duedal S, Kirner J, McGuffog L, Last J, et al. 2005. Cancer risks and mortality in heterozygous ATM mutation carriers. *J. Natl. Cancer Inst.* 97:813–22
114. Thompson D, Easton D. 2001. Variation in cancer risks, by mutation position, in BRCA2 mutation carriers. *Am. J. Hum. Genet.* 68:410–19
115. Thompson D, Easton D. 2002. Variation in *BRCA1* cancer risks by mutation position. *Cancer Epidemiol. Biomark. Prev.* 11:329–36
116. Thompson D, Easton DF. 2002. Cancer incidence in BRCA1 mutation carriers. *J. Natl. Cancer Inst.* 94:1358–65
117. Thompson D, Seal S, Schutte M, McGuffog L, Barfoot R, et al. 2006. A multicenter study of cancer incidence in CHEK2 1100delC mutation carriers. *Cancer Epidemiol. Biomark. Prev.* 15:2542–45
118. Tommiska J, Seal S, Renwick A, Barfoot R, Baskcomb L, et al. 2006. Evaluation of RAD50 in familial breast cancer predisposition. *Int. J. Cancer* 118:2911–16
119. Walsh T, Casadei S, Coats KH, Swisher E, Stray SM, et al. 2006. Spectrum of mutations in *BRCA1*, *BRCA2*, *CHEK2*, and *TP53* in families at high risk of breast cancer. *JAMA* 295:1379–88
120. Ware MD, DeSilva D, Sinilnikova OM, Stoppa-Lyonnet D, Tavtigian SV, Mazoyer S. 2006. Does nonsense-mediated mRNA decay explain the ovarian cancer cluster region of the *BRCA2* gene? *Oncogene* 25:323–28
121. Williams WR, Anderson DE. 1984. Genetic epidemiology of breast cancer: segregation analysis of 200 Danish pedigrees. *Genet. Epidemiol.* 1:7–20
122. Wooster R, Bignell G, Lancaster J, Swift S, Seal S, et al. 1995. Identification of the breast cancer susceptibility gene *BRCA2*. *Nature* 378:789–92
123. Wooster R, Neuhausen SL, Mangion J, Quirk Y, Ford D, et al. 1994. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12–13. *Science* 265:2088–90
124. Xia B, Dorsman JC, Ameziane N, de Vries Y, Rooimans MA, et al. 2007. Fanconi anemia is associated with a defect in the BRCA2 partner PALB2. *Nat. Genet.* 39:159–61
125. Xia B, Sheng Q, Nakanishi K, Ohashi A, Wu J, et al. 2006. Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. *Mol. Cell* 22:719–29